

(12) DEMANDE INTERNATIONALE PUBLIÉE EN VERTU DU TRAITE DE COOPÉRATION
EN MATIÈRE DE BREVETS (PCT)

(19) Organisation Mondiale de la Propriété
Intellectuelle
Bureau international



(43) Date de la publication internationale
5 décembre 2002 (05.12.2002)

PCT

(10) Numéro de publication internationale
WO 02/097100 A2

(51) Classification internationale des brevets⁷ : C12N 15/79

(21) Numéro de la demande internationale :
PCT/FR02/01782

(22) Date de dépôt international : 28 mai 2002 (28.05.2002)

(25) Langue de dépôt : français

(26) Langue de publication : français

(30) Données relatives à la priorité :
0106921 28 mai 2001 (28.05.2001) FR

(71) Déposant (pour tous les États désignés sauf US) :
GENOWAY [FR/FR]; 181, Avenue Jean-Jaurès, Immeuble
Château Briand, F-69007 Lyon (FR).

(72) Inventeurs; et

(75) Inventeurs/Déposants (pour US seulement) : LAPIZE-
GAUTHEY, Christine [FR/FR]; 4, rue Latreille, F-38200
Vienne (FR). FRAICHARD, Alexandre [FR/FR]; 144,
avenue des États-Unis, F-78000 Versailles (FR).

(74) Mandataires : MARTIN, Jean-Jacques etc.; Cabinet
Regimbeau, 26, rue de Chazelles, F-75847 Paris Cedex 17
(FR).

(81) États désignés (national) : AE, AG, AL, AM, AT, AU, AZ,
BA, BB, BG, BR, BY, BZ, CA, CH, CN, CO, CR, CU, CZ,
DE, DK, DM, DZ, EC, EE, ES, FI, GB, GD, GE, GH, GM,
HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK,
LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX,
MZ, NO, NZ, OM, PH, PL, PT, RO, RU, SD, SE, SG, SI,
SK, SL, TJ, TM, TR, TT, TZ, UA, UG, US, UZ, VN,
YU, ZA, ZM, ZW.

(84) États désignés (régional) : brevet ARIPO (GH, GM, KE,
LS, MW, MZ, SD, SL, SZ, TZ, UG, ZM, ZW), brevet
eurasien (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), brevet
européen (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR,
IE, IT, LU, MC, NL, PT, SE, TR), brevet OAPI (BF, BJ,
CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN,
TD, TG).

Déclaration en vertu de la règle 4.17 :

— relative à la qualité d'inventeur (règle 4.17.iv) pour US
seulement

Publiée :

— sans rapport de recherche internationale, sera republiée
dès réception de ce rapport

En ce qui concerne les codes à deux lettres et autres abrégia-
tions, se référer aux "Notes explicatives relatives aux codes et
abrégiactions" figurant au début de chaque numéro ordinaire de
la Gazette du PCT.

(54) Title: CLONING VECTORS FOR HOMOLOGOUS RECOMBINATION AND METHOD USING SAME

(54) Titre : VECTEURS DE CLONAGE POUR RECOMBINAISON HOMOLOGUE ET PROCÉDE LES UTILISANT

(57) Abstract: The invention concerns novel vectors for use both in creating genomic DNA banks and as vectors for carrying out homologous recombination reactions in host cells, in particular for improving selection of said homologous recombination events, and decreasing the time for obtaining a final vector for performing the homologous recombination reaction. The invention also concerns a method using said vectors.

(57) Abrégé : La présente invention se rapporte à des nouveaux vecteurs utilisables à la fois pour créer des banques d'ADN gé-
nomiques et en tant que vecteurs pour la réalisation de réactions de recombinaison homologue dans des cellules hôtes, permettant
notamment d'améliorer la sélection desdits événements de recombinaison homologue, et de diminuer le temps d'obtention du vec-
teur final pour effectuer la réaction de recombinaison homologue. L'invention a également pour objet un procédé mettant en oeuvre
ces vecteurs.

WO 02/097100 A2

VECTEURS DE CLONAGE POUR RECOMBINAISON HOMOLOGUE ET PROCÉDE LES UTILISANT

La présente invention se rapporte à de nouveaux vecteurs utilisables à la fois
5 pour créer des banques d'ADN génomiques et en tant que vecteurs pour la
réalisation de réactions de recombinaison homologue dans des cellules hôtes,
permettant notamment d'améliorer la sélection desdits événements de
recombinaison homologue et de diminuer le temps d'obtention du vecteur final pour
effectuer la réaction de recombinaison homologue. L'invention a également pour
10 objet un procédé mettant en œuvre ces vecteurs. Ces nouveaux vecteurs et le
procédé les utilisant sont spécialement adaptés pour la recombinaison homologue
dans des cellules souches de mammifères.

L'obtention de souris porteuses de modifications génétiques programmées
par l'expérimentateur a entièrement modifié l'étude de quasiment tous les aspects
15 de la biologie de cet animal et de ces différents systèmes (immunitaire, nerveux,
hématopoïétique, etc.). En outre, ces méthodes de modification du patrimoine
génétique de la souris permettent la création de modèles murins de maladies
génétiques humaines qui sont précieux pour l'étude de leur physiopathologie et
éventuellement la mise au point de thérapeutiques appropriées. Cela a été rendu
20 possible grâce à la rencontre heureuse de deux approches très différentes : l'une a
abouti à l'isolement *in vitro* de cellules souches embryonnaires (cellules ES) et
l'autre a permis d'identifier, chez les cellules d'eucaryotes supérieurs, les conditions
nécessaires au processus de recombinaison homologue entre un ADN exogène
connu et la séquence homologue dans le chromosome.

25

La recombinaison homologue dans les cellules ES : création de mutations nulles (souris knock-out)

Des études effectuées dans les années 1980 (Smithies et Capecchi par
exemple) ont démontré que les cellules embryonnaires souches (ES) possèdent
30 l'appareil enzymatique nécessaire à la recombinaison homologue (RH) entre une
séquence d'ADN exogène et la séquence homologue présente dans le génome de la
souris, même s'il s'agit d'un événement rare comparativement à l'intégration
aléatoire de ce même fragment d'ADN (Smithies et al., 1985 ; Wong and Capecchi,

1986). En dépit de cette rare fréquence d'apparition, la technique de recombinaison homologue a pu être étendue à l'étude de nombreux gènes (Jackson laboratory) grâce à différentes astuces de sélection et de criblage.

5 Le Knock-in : l'autre aspect de la recombinaison homologue.

Une variante intéressante des vecteurs de ciblage pour l'obtention de mutations nulles résulte de l'introduction, en phase avec la séquence codante du gène cible, d'un ADNc d'intérêt.

Après la recombinaison homologue, l'ADNc inséré dans le gène cible est
10 exprimé sous le contrôle du promoteur endogène. Le choix de l'ADNc dépend du projet et du but recherché. Par exemple, il peut s'agir de la séquence codante d'un gène rapporteur comme la β -galactosidase. L'expression de cette protéine mime alors l'expression du gène ciblé, ce qui permet de déterminer son profil d'expression (Li et al., 1997) et/ou de suivre le devenir des cellules exprimant le
15 gène ciblé (Schneider-Maunoury et al., 1993 ; Tajbakhsh et al., 1996).

Des mutations nulles aux mutations subtiles

Si les mutations nulles (knock-out) représentent un instrument d'analyse génétique très puissant, il est clair que d'autres types de mutations plus fines
20 (mutations ponctuelles, petites délétions ou insertions) sont très utiles pour affiner les modèles d'études. Lorsque des allèles mutés sont créés, il est particulièrement nécessaire d'éliminer les séquences de sélection qui pourraient interférer avec la régulation de l'expression du gène cible ou de gènes adjacents (Fiering et al., 1999). Plusieurs stratégies peuvent être utilisées pour créer des mutations "propres"
25 (Moore et al., 1998 ; Cohen-Tannoudji et Babinet, 1998) comme le "Plug and socket" (Detloff et al., 1994), le "hit and run" (Hasty et al., 1991) ou le système de recombinaisons Cre/loxP ou FLP/FRT (Kilby et al., 1993).

La stratégie des recombinaisons (Cre/loxP, FRT/FLP...) repose sur l'utilisation d'une protéine, la recombinaise, et de séquences cibles. Dans le système
30 Cre/loxP, la protéine Cre est une recombinaise, identifiée chez le bactériophage P1, qui agit lorsqu'elle reconnaît dans un segment d'ADN, une séquence de 34 pb appelée loxP (Sauer and Henderson, 1988). Après l'excision par la Cre du fragment se trouvant entre les deux séquences loxP, l'allèle muté porte un site loxP. Aucune

indication d'une interférence de ce site avec l'expression génétique n'a pu être mis en évidence.

La mutagenèse conditionnelle

5 La stratégie Cre/loxP (ou FRT/FLP, qui est similaire) permet l'apparition conditionnelle d'une mutation chez l'animal en cours de développement ou chez l'adulte (Cohen-Tannoudji et Babinet, 1998 ; Gu et al., 1994). Il s'agit dans un premier temps de créer des souris porteuses d'un allèle flanqué de deux séquences loxP encadrant une partie essentielle d'un gène d'intérêt sans pour autant altérer son
10 fonctionnement, en les plaçant par exemple dans des introns (gène "floxé"). On accouple une souris ainsi produite avec une autre souris transgénique exprimant la recombinaise dans un type cellulaire particulier. Cette stratégie est potentiellement très efficace, car elle permet non seulement de contourner le problème de la létalité embryonnaire qui se produit lorsque toutes les cellules de l'embryon portent la
15 mutation, mais aussi d'examiner l'effet de cette mutation dans n'importe quel tissu pourvu que l'on dispose d'une lignée de souris transgéniques exprimant la protéine Cre dans le tissu en question (Xu et al., 1999 ; Shibata et al., 1997 ; Kulkarni et al., 1999 ; Tsien et al., 1996 ; Harada et al., 1999).

 Un raffinement supplémentaire consiste à contrôler l'induction de la
20 mutation, non seulement dans l'espace mais également dans le temps grâce à des systèmes inductibles. Dans cette optique, la protéine Cre est exprimée sous la forme d'une protéine de fusion avec un domaine de fixation à un ligand. Cette protéine de fusion ne présente pas d'activité Cre. En revanche, en présence du ligand approprié, un changement de conformation permet la restauration de l'activité de Cre. Ainsi
25 chez les souris transgéniques portant les deux allèles du gène d'intérêt "floxés" et exprimant la protéine Cre fusion (récepteur au ligand associé à la protéine) dans un tissu particulier (Shibata et al., 1997), l'administration du ligand approprié provoquera l'induction de la mutation nulle dans ce type de tissu au moment voulu. Il est également possible de contrôler l'expression de la recombinaise en utilisant des
30 systèmes qui permettent d'induire ou de réprimer la transcription d'un gène rapporteur. Le système le plus documenté utilise les propriétés du couple opérateur/répresseur de l'opéron bactérien tétracycline (tet) (Baron U. et al, 1999)

Toutes ces techniques prometteuses présentent cependant deux inconvénients majeurs, la construction de vecteurs complexes et le temps indispensable pour réaliser et analyser ces modèles murins.

5 Ainsi, on estime qu'aujourd'hui, le temps d'obtention d'une souris knock-in ou knock-out conditionnel est d'environ 15 à 18 mois. En moyenne, il faut en effet environ 8 à 9 mois pour cloner, cartographier, et construire le vecteur cible, 1 à 2 mois pour la culture de cellules ES et enfin 6 à 8 mois d'animalerie et de croisement.

10 L'objet de la présente invention est de réduire les délais de manipulations de biologie moléculaire qui devraient, par l'utilisation des vecteurs et/ou procédés selon l'invention, pouvoir être menées en 3 mois au lieu de 8 ou 9 mois.

15 Ce gain de temps est significatif tout en présentant un risque plus faible de non-réussite grâce à l'élimination des étapes de sous-clonages et à la diminution du nombre d'étapes de clonage, grâce à l'utilisation des vecteurs selon l'invention. Il est estimé que l'utilisation de ces vecteurs permet de descendre en dessous de la barre des 12 mois pour produire un modèle murin.

20 L'objectif de la présente invention est de fournir un nouveau vecteur de base, adapté à la création d'une banque d'ADN génomique, ledit vecteur étant également adapté pour pouvoir être transformé de façon directe (sans sous-clonage) afin d'effectuer les réactions de recombinaison homologue dans l'hôte cible. L'invention divulgue ainsi un procédé permettant d'accélérer la fabrication des vecteurs utilisés pour la mise en œuvre de la recombinaison homologue, dans tout organisme, et de préférence dans des cellules eucaryotes, notamment des cellules souches ES issues d'organismes pluricellulaires.

25

Ainsi, dans un premier aspect, la présente invention a pour objet un vecteur de clonage adapté pour la création d'une banque d'ADN génomique, caractérisé en ce qu'il présente une cassette comprenant :

30 - un polylinker permettant l'intégration d'un fragment d'ADN génomique, ledit polylinker étant flanqué de sites rares de restriction, et que ladite cassette est flanquée de sites très rares de restrictions.

Un élément est « flanqué de sites de restriction » lorsqu'un site de restriction au moins est présent à chaque extrémité de l'élément.

Dans un mode de réalisation préférée, ladite cassette comprend également au moins un gène de sélection négative.

Le vecteur selon l'invention est parfaitement adapté pour la création de banque d'ADN génomique, et permet notamment le clonage de fragments d'ADN de taille moyenne voire de grande taille. L'homme du métier sait ce que signifie qu'un vecteur est adapté à la création d'une banque d'ADN, notamment génomique, c'est-à-dire qu'il permet d'insérer et d'obtenir la stabilité des séquences d'ADN insérées, notamment pour ce qui est du maintien du vecteur dans la cellule hôte, en particulier lors de la réplication, et pour lequel on observe un taux faible d'événements de recombinaison du fragment d'ADN avec lui-même et/ou perte de séquences d'ADN exogène. Des exemples de vecteurs utilisables sont donnés ci-après.

Par « fragments de taille moyenne », on entend notamment des fragments d'une taille supérieure à 15 kilobases (kb), de préférence comprise entre 15 kb et 75 kb, de préférence comprise entre environ 17 kb et environ 50 kb, de préférence comprise entre environ 20 kb et environ 30-35 kb.

Par « sites rares de restriction », on entend des sites de restriction dont la fréquence de coupure est supérieure à 10 kb, de préférence 15 kb, de façon plus préférée 20 kb dans le génome humain. Parmi les enzymes rares, on peut notamment citer *PmeI*, *SgrAI*, *RsrII*, *ClaI*, *NotI*, *PacI*, *SrfI*, *NheI*, *FseI*, *NsiI*.

Un site de restriction est dit « très rare » lorsque sa fréquence de coupure dans le génome humain est supérieure à 100, de préférence 200 kb (par exemple *AscI*). Parmi les sites de restriction très rares utilisables, on peut citer les « Homing endonucleases ». Ces enzymes sont des protéines codées par des gènes possédant des introns s'auto-épissant. Ces enzymes effectuent des coupures site-spécifiques dans de l'ADN double brin, et reconnaissent généralement des sites de 18-20 bases ou plus. On note en particulier *I-PpoI*, *I-CeuI*, *PI-PsI*, *I-SceI*, *PI-SceI*.

Par ailleurs, ainsi qu'il sera vu, le vecteur de clonage selon la présente invention, et en particulier lorsqu'il présente au moins un gène de sélection négative, est également particulièrement adapté pour la recombinaison homologue dans des cellules de mammifères, notamment les cellules souches murines, de lapin, porcines, bovines, humaines.

Dans un cas particulier, le vecteur selon l'invention présente, dans son polylinker pour le clonage de l'ADN exogène, un maximum de 4, de préférence 3, de préférence deux sites de restrictions. Il est particulièrement avantageux de réduire le nombre de sites de restrictions situés dans le polylinker utilisé pour l'insertion du fragment d'ADN exogène, ou situés dans le squelette du vecteur, afin de pouvoir utiliser les enzymes de restriction correspondantes après clonage du fragment d'ADN exogène, pour vérifier notamment la nature dudit fragment, voire son orientation, ainsi que la nature, l'orientation et l'emplacement exact des modifications introduites dans le fragment d'ADN exogène. Ainsi, il est possible que le polylinker ne présente qu'un seul site de restriction.

Il existe aujourd'hui différents types de vecteurs permettant la construction de banques d'ADN. On peut notamment citer les cosmides ou les chromosomes artificiels. Ces derniers, qui peuvent être adaptés aux bactéries (BAC), aux levures (YAC) ou aux cellules mammifères (MAC) permettent l'intégration stable de fragments d'ADN pouvant aller jusqu'à 200 kb ou plus. Les cosmides ont une capacité plus limitée (environ 20 à 40 kb), mais qui est toutefois plus importante que les plasmides classiques, par exemple ceux dérivés de pBlueScript qui ne peuvent recevoir au maximum qu'environ 14-15 kb.

Les chromosomes artificiels de bactéries sont des vecteurs particulièrement attractifs, car ils peuvent être utilisés dans des hôtes faciles à manipuler (notamment *Escherichia coli*), et assurent un maintien stable des fragments d'ADN introduits.

Ainsi, dans un cas particulier de l'invention, le vecteur est dérivé d'un cosmide, ou d'un vecteur de chromosome artificiel, de préférence de bactérie.

Par « vecteur dérivé d'un autre vecteur », on entend que le vecteur final possède globalement le même squelette (notamment les origines de répllication...) que le vecteur de départ dont il est dérivé (vecteur parent). Il est important de noter que ceci signifie également que, de préférence, les éléments intergéniques sont conservés. Les modifications apportées dans le vecteur de base sont de telles sortes que l'hôte de répllication du vecteur reste le même, et ne modifient pas les propriétés fondamentales du vecteur (nombre de copies, taille de l'insert, stabilité...). Il est à noter que l'on peut envisager de modifier le gène de sélection du vecteur dérivé par rapport au vecteur parent, pour autant que les propriétés de réception d'ADN exogène du vecteur dérivé ne soient pas altérées par rapport au vecteur parent.

On choisit de préférence un vecteur dérivé du corps du vecteur de chromosome artificiel de bactérie pBeloBAC11, bien connu de l'homme du métier, dont la séquence peut notamment être obtenue dans GenBank sous le numéro d'accès U51113.

5 Le vecteur selon l'invention est préférentiellement basé sur le squelette de pBeloBAC11, digéré par *SalI*, et reliqué. Ainsi, le vecteur selon l'invention est dépourvu des séquences *cosN*, *loxP*, et *lacZ*, initialement présentes dans pBeloBAC11. Une telle modification permet en effet de supprimer certaines séquences d'origine bactérienne (*lacZ*), ou utiles dans la levure (*cosN*), ou pouvant
10 interférer ultérieurement avec une recombinaison utile pour le procédé selon l'invention (*loxP*). La suppression des éléments d'origine bactérienne ou de levure est en effet avantageuse pour améliorer la fréquence des événements de recombinaison dans les cellules de mammifères. Le maintien du site *loxP* pourrait s'avérer gênant pour l'utilisation ultérieure d'une recombinaison Cre sur les vecteurs
15 obtenus.

Dans un cas particulier de l'invention, on a supprimé, dans le vecteur selon l'invention, au moins un site de restriction choisi parmi les sites présents 1, 2 ou 3 fois dans le squelette du vecteur parent.

L'absence d'au moins un des sites de restrictions précités dans le squelette
20 du vecteur selon l'invention permet notamment d'utiliser ce site pour l'étude des fragments intégrés dans ledit vecteur, ou pour introduire par procédé enzymatique les gènes de sélection modifiant le fragment d'ADN exogène avant son utilisation pour la recombinaison homologue proprement dite.

La modification du squelette du vecteur pour supprimer les sites précités
25 peut être effectuée par mutagenèse dirigée, afin que l'enzyme de restriction correspondante ne coupe plus le squelette du vecteur.

D'une façon préférée, et notamment lorsque le vecteur selon l'invention est dérivé du vecteur pBeloBAC11, on a supprimé, dans son squelette au moins un site de restriction choisi parmi *ApaI*, *BstEII*, *SacII*, *SfiI*, *SpeI*, *SphI*, *StuI*, *XhoI*, *BssHII*,
30 *EcoRI*, *EcoRV*, *KpnI*, *NdeI*, *NotI*, *NruI*, *PvuI*, *SgrAI*, *XbaI*, *PstI*, *SalI* et *SmaI*. Ainsi que précisé plus haut, il peut être avantageux de supprimer le maximum de ces sites de restriction, c'est-à-dire au minimum 2 de ces sites, au mieux 3, 4 ou 5, dans le

meilleur des cas, de n'en garder aucun. Dans un cas particulier, le squelette du vecteur selon l'invention ne possède plus le site *BstEII*.

Dans un mode de réalisation particulier, la cassette du vecteur selon l'invention présente un site très rare de restriction à l'une de ses extrémités et deux
5 sites très rares de restriction à l'autre extrémité, l'un de ces deux sites étant de préférence identique au site situé à la première extrémité de la cassette.

La présence de ces sites permet soit de linéariser le vecteur selon l'invention (utilisation du site présent à une seule extrémité de la cassette), soit d'exciser l'insert intégré dans le vecteur (utilisation du site situé aux deux extrémités de la
10 cassette).

Dans un cas particulier et préféré, le polylinker du vecteur selon l'invention permettant l'intégration de l'ADN exogène est flanqué de plusieurs (notamment un nombre égal ou supérieur à 2, 3, ou 4) sites rares de restrictions choisis notamment parmi *PmeI*, *SgrAI*, *RsrII*, *ClaI*, *NotI*, *PacI*, *SrfI*, *NheI*, *FseI*.

15 Le vecteur selon l'invention contient ainsi au moins un gène de sélection négative. On peut notamment choisir des gènes suicides, c'est-à-dire qui codent pour des produits capables de transformer une substance inactive en substance cytotoxique. On cite notamment les gènes de la thymidine kinase du virus HSV-1, utilisable avec le ganciclovir ou l'acyclovir, le gène DTA, codant pour le fragment
20 A de la toxine diphtérique, décrit dans Yagi *et al.*, (1990, PNAS, 87, 9918-22), toxique par sa seule présence. Il existe aujourd'hui plusieurs gènes suicides à la disposition de l'homme du métier, pouvant être utilisés comme marqueurs de sélection négative. On peut aussi citer le cytochrome p450 de rat et la cyclophosphamide (Wei *et al.*, 1994, Human Gene Ther. 5, 969-978), la purine
25 nucleoside phosphorylase d'*Escherichia coli* (*E. coli*) et la 6-méthylpurine deoxyribonucleoside (Sorscher *et al.*, 1994, Gene Therapy 1, 223-8), la guanine phosphoribosyl transférase d'*E. coli* et la 6-thioxanthine (Mzoz *et al.*, 1993, Human Gene Ther. 4, 589-595), les cytosine deaminase (CDase) ou uracil phosphoribosyl transférase (UPRTase) qui peuvent être utilisées avec la 5-fluorocytosine (5-FC).

30 Le gène de sélection négative est ainsi situé à une extrémité du polylinker utilisé pour le clonage du fragment d'ADN génomique, mais le vecteur selon l'invention contient de préférence deux gènes de sélection négative (éventuellement identiques), chacun étant situé de part et d'autre dudit polylinker.

D'une façon préférée, le vecteur de clonage selon l'invention contient deux copies du même gène de sélection négative, flanquant le polylinker utilisé pour le clonage du fragment d'ADN génomique, lui-même flanqué de sites rares de restriction (figure 1).

- 5 On obtient donc une construction ainsi faite (de 5' vers 3') :
- au moins un site très rare de restriction,
 - un marqueur de sélection négative,
 - une cassette SR1 possédant au moins un (de préférence plusieurs) site(s) rare(s) de restriction,
 - 10 - le polylinker présentant un nombre réduit de sites de restriction,
 - une cassette SR2 contenant au moins un (de préférence plusieurs) site(s) rare(s) de restriction, de préférence différent(s) du (des) site(s) de la première cassette,
 - un marqueur de sélection négative, éventuellement identique au premier
 - 15 marqueur,
 - au moins un site très rare de restriction, de préférence deux sites très rares de restriction, l'un étant de préférence identique au premier site mentionné ci-dessus, l'autre différent.

La présence de plusieurs sites très rares de restriction permet en effet de
20 linéariser le vecteur de clonage selon l'invention ou d'isoler l'insert situé entre deux sites très rares de restriction.

Il est à noter que, de préférence, les sites rares de restriction contenus dans une des cassettes sont tous différents de ceux contenus dans l'autre cassette. Ceci permet en effet de linéariser le vecteur selon l'invention après clonage du fragment
25 d'ADN génomique, et, dans la mesure où il est intéressant, pour des événements de recombinaison homologue dans les cellules mammifères, de disposer d'un bras long d'homologie et d'un bras court d'homologie, de pouvoir utiliser l'exonucléase III afin de réduire l'un des deux bras. Il est à noter que les cassettes SR1 et SR2 peuvent aussi contenir des sites de restriction moins rares, tels notamment *SacI*,
30 *SwaI*, *SphI* ou *SalI*.

Le vecteur selon l'invention permet donc la construction d'une banque d'ADN génomique de l'organisme dans lequel on désire effectuer la recombinaison homologue. De préférence, cette banque est composée de fragments de taille

moyenne, afin de pouvoir effectuer les modifications nécessaires à la recombinaison homologue directement dans le vecteur ayant servi à la construction de la banque, sans qu'il soit nécessaire de sous-cloner le fragment d'ADN génomique portant le locus cible, après son identification dans la banque. Il est toutefois important de
5 noter que la taille des fragments dans la banque peut être proportionnelle à la taille du génome de l'organisme cible, dans la mesure où il est intéressant (notamment pour des raisons pratiques de manipulation) de limiter le nombre de clones présents dans la banque. Ainsi, pour un génome tel que celui de la souris, on choisit de fabriquer une banque de fragments d'une taille moyenne égale à 20-30 kb.

10 Dans un mode de réalisation préféré, le vecteur selon l'invention est le vecteur représenté par SEQ ID N° 1, possédant en outre les caractéristiques suivantes : 13478 pb, circulaire, nt 4629..5597 protéine SOPB, nt 3454..4626 protéine SOPA, nt 2120..2872 protéine REP, nt 999..1343 protéine Resolvase, nt 114..780 gène de résistance au Chloramphénicol, nt 6443..9851 gène DTA, nt
15 10037..13451 gène DTA, nt 9943..9962 oligonucléotide Sp6, nt 9902..9922 oligonucléotide T7.

La présente invention concerne donc également une banque d'ADN génomique de fragments de taille moyenne, fabriquée dans un vecteur selon l'invention. L'homme du métier connaît les différents moyens de fabriquer une
20 banque d'ADN génomique dans un vecteur approprié, notamment par digestion partielle de l'ADN génomique.

Le vecteur selon l'invention a de plus été optimisé afin qu'il soit parfaitement adapté à un procédé de mise en œuvre de recombinaison ciblée dans le génome d'un organisme cible, à partir d'une banque de fragments génomiques qui y
25 sont clonés.

Ainsi, on peut introduire les éléments modifiant le fragment portant le locus cible identifié dans la banque d'ADN génomique directement dans le vecteur le portant et effectuer la réaction de recombinaison homologue dans l'hôte cible, sans avoir besoin de sous-cloner le locus cible.

30 En effet, le fait que le vecteur selon l'invention porte un gène de sélection négative permet de sélectionner plus facilement les événements de recombinaison homologue.

Après sélection du clone portant le locus cible dans la banque d'ADN génomique, on peut introduire le ou les fragments d'ADN qui seront utilisés pour modifier le gène cible *in vivo* (par exemple les gènes de sélection positive ou sites de reconnaissance de la recombinaise Cre). Ces fragments d'ADN peuvent être
5 introduits dans le vecteur par clonage classique (utilisation de sites enzymatiques) ou par recombinaison homologue dans les bactéries, ou encore par l'utilisation de transposons ou par toute autre technique adaptée. L'utilisation de transposons est préférée, car elle permet de réduire le temps nécessaire à l'obtention du locus modifié, puisqu'il n'est pas besoin d'étudier préalablement la cartographie dudit
10 locus.

Ainsi, la présente invention a également pour objet un procédé pour effectuer une recombinaison ciblée dans un organisme, à partir d'une banque d'ADN génomique dans un vecteur selon l'invention comprenant :

- a) l'incubation d'un vecteur de ladite banque génomique avec au moins un
15 fragment d'ADN comprenant au moins un transposon, ledit transposon comprenant
 - un gène de sélection positive,
 - éventuellement un gène de sélection négative et/ou un gène marqueur,
20 lesdits gènes de sélection étant éventuellement encadrés par les sites d'action I d'une recombinaise site-spécifique I, et de façon optionnelle
 - au moins un site d'action II d'une recombinaise site-spécifique II différente de la première recombinaise I,
25 ledit site d'action II n'étant pas situé entre lesdits sites d'action I de la première recombinaise I,
en présence d'une enzyme ayant une activité de transposase pour le(s)dit(s) transposon(s), de telle sorte que le(s)dit(s) transposon(s) est (sont) transféré(s) dans ledit fragment génomique,
- b) l'introduction dudit vecteur, éventuellement linéarisé, portant ledit
30 fragment génomique dans lequel est (sont) inséré(s) le(s)dit(s) transposons dans une cellule hôte cible issue dudit organisme,
- c) la sélection des événements de recombinaison homologue dans ladite cellule cible par l'utilisation du (des) gène(s) de sélection positive

porté(s) par le(s)dit(s) transposon(s), et éventuellement du gène de sélection négative porté par ledit vecteur de clonage.

De préférence, ledit fragment génomique modifié est un fragment de taille moyenne.

- 5 Le procédé selon l'invention permet l'intégration ciblée, et la recombinaison homologue dans un organisme cible, de préférence un organisme mammifère, de façon plus préférée des cellules de rongeurs (notamment murines), de lapins, porcines, ovines, bovines, voire humaines.

- 10 Afin de vérifier les loci d'intégration des transposons dans les fragments génomiques portés par les vecteurs selon l'invention, on peut notamment utiliser les enzymes de restriction coupant aux sites rares dans les cassettes SR1 et SR2. On utilise aussi éventuellement ces enzymes pour linéariser le vecteur, et former les petit et long bras d'homologie grâce à l'action ciblée de l'exonucléase III.

- 15 Le procédé selon l'invention est avantageusement complété en faisant suivre l'étape c) d'une étape d) :

d) la soumission des cellules transformées par recombinaison homologue et sélectionnées dans l'étape c) à l'action de la recombinaise I, afin d'éliminer le(s) gène(s) de sélection positive et éventuellement négative porté(s) par le(s) transposon(s).

- 20 De façon préférée, l'organisme cible est donc un organisme pluricellulaire, et ladite cellule cible issue dudit organisme est une cellule souche.

- 25 Dans un cas préféré, on induit la délétion d'un fragment génomique et/ou l'insertion d'un fragment exogène dans ledit organisme après l'étape c) ou éventuellement l'étape d) par la mise en contact des cellules sélectionnées dans l'étape c) ou dérivées desdites cellules avec la recombinaise II. Ceci permet ainsi d'obtenir une inactivation conditionnelle du gène que l'on a ciblé, notamment lorsque la construction a été réalisée de telle façon que les sites loxP se retrouvent dans des introns du gène que l'on désire inactiver. Cette inactivation conditionnelle est effectuée *in vivo*, ainsi que décrit dans l'introduction.

- 30 La présente invention concerne également un kit comprenant :

- un vecteur de clonage selon l'invention,
- au moins un fragment d'ADN comprenant au moins un transposon, ledit transposon comprenant un gène de sélection positive, et éventuellement

un gène de sélection négative et/ou un gène marqueur, le(s)dit(s) gène(s) de sélection étant éventuellement, et de façon préférée, encadré(s) par les sites d'action I d'une recombinaise site-spécifique I.

De façon optionnelle, ledit transposon comprend également un gène de sélection négative ou un gène marqueur situé à côté du gène de sélection positive, situé ou non entre les sites d'action I de ladite recombinaise site-spécifique I lorsque celle-ci est présente.

Le transposon utilisable pour la présente invention peut être de toute sorte, notamment le transposon Tn5. Le transposon Tn5 est choisi de préférence pour son intégration unidirectionnelle dans l'ADN cible (de 5' vers 3'), pour une utilisation directe avec une transposase sans avoir à rajouter de cofacteurs et pour sa commercialisation en kit.

Pour contourner le problème d'instabilité, s'il se présente, l'utilisation de différents types de transposons (par exemple Tn5, Tn10, Tn7) peut être envisagée (Brune et al., 1999; Chatterjee and Coren, 1997; Goryshin et al., 2000; Goryshin and Reznikoff, 1998; Stellwagen and Craig, 1997; Westphal and Leder, 1997; Yang et al., 1997).

Certains gènes de sélection négative utilisables pour la mise en œuvre de la présente invention ont été cités plus haut. Parmi les gènes marqueurs utilisables, on peut citer le gène lacZ, ou les gènes codant pour des protéines fluorescentes (FP).

Les gènes de sélection positive sont bien connus de l'homme du métier et sont préférentiellement des gènes de résistance à un antibiotique, tel les gènes de résistance à la kanamycine, qui permet également la résistance à la néomycine dans les cellules mammifères, résistance à l'hygromycine, à la zéocine, à la blasticidine...

De façon optionnelle, le transposon comprend également au moins un site d'action II d'une recombinaise site-spécifique II différente de la première recombinaise I, le(s)dit(s) site(s) d'action II n'étant pas situé(s) entre lesdits sites d'action I de la première recombinaise I.

De façon préférée, le kit selon l'invention comprend deux fragments d'ADN, chacun comprenant un transposon, chaque transposon portant un gène de sélection positive différent, et éventuellement un gène de sélection négative, le gène de sélection négative étant alors de préférence identique pour les deux transposons,

lesdits gènes de sélection positive et négative étant de préférence encadrés par les sites d'action d'une recombinaise I.

Dans un mode de mise en œuvre préférée de l'invention, chaque transposon comprend une séquence correspondant à un site d'action d'une recombinaise II
5 différente de la première recombinaise I.

Ainsi, dans un cas préféré, ladite recombinaise I est la recombinaise FLP, ladite recombinaise II étant la recombinaise Cre. Il est alors avantageux que l'un des transposons contiennent les séquences FRT natives reconnues par la recombinaise FLP, l'autre transposon contenant alors les séquences FRT* mutées, telles que
10 décrites par Schalke et Bode (1994).

Afin de mettre en œuvre le procédé selon l'invention, décrit plus haut, pour faciliter les événements de recombinaison ciblée, il est avantageux que le kit selon l'invention comprenne également une enzyme ayant une activité de transposase pour le(s)dit(s) transposon(s), ainsi que les instructions de mise en œuvre pour
15 l'insertion du(des)dit(s) transposon(s) dans le fragment génomique situé dans ledit vecteur de clonage.

DESCRIPTION DES FIGURES

Figure 1 : carte schématique d'un vecteur modifié selon l'invention. SR1 et SR2 :
20 cassettes présentant les sites de restriction rares. PL : polylinker pour le clonage des fragments génomiques. DTA : gène de la sous-unité A de la toxine diphtérique.

Figure 2 : description de l'insertion de deux transposons dans un fragment d'ADN dans le vecteur selon l'invention. DTA : gène du fragment A de la toxine diphtérique, présent dans le vecteur selon l'invention. SR1 et SR2, cassettes
25 contenant au moins un site de restriction rare. Site très rare : site notamment d'Homing endonucléase. Kana : gène de résistance à la kanamycine en bactéries, à la néomycine dans les cellules eucaryotes. Zeocin : gène de résistance à la zéocine. TK : gène de la thymidine kinase du HSV1. FRT : sites de reconnaissance de la recombinaise FLP. FRT* : sites mutés de reconnaissance de la recombinaise FLP.
30 LoxP : sites de reconnaissance de la recombinaise Cre. OE/S : séquences permettant de déterminer le locus d'intégration du transposon.

Figure 3 : Schéma de la recombinaison homologue d'un vecteur selon l'invention après intégration de transposons dans un locus génomique. Les boîtes noires 1, 2, 3,

4 représentent les exons du gène cible. Les autres sigles ont la signification précédente. L'événement de recombinaison homologue est sélectionné par la résistance à la néomycine (présence du gène kanamycine) et à la Zéocine, et l'absence des gènes DTA du vecteur (sélection négative).

- 5 Figure 4 : Action de la recombinase FLP dans les cellules sélectionnées pour la recombinaison homologue pour obtenir un locus final knock-out conditionnel en supprimant les séquences d'ADN exogènes (gènes de résistance).

- Figure 5 : Action de la recombinase Cre sur les cellules sélectionnées pour la recombinaison homologue pour obtenir un locus final knock-out, et production
10 d'une protéine tronquée.

Figure 6 : Action de la recombinase Cre *in vivo* dans l'animal transgénique dérivant des cellules ayant été soumises à la recombinase FLP pour obtenir un locus final knock-out, et production d'une protéine tronquée de manière conditionnelle.

- Figure 7 : carte du plasmide pBeloBAC11, utilisé en tant que vecteur parent pour
15 certains vecteurs selon l'invention.

- Figure 8 : stratégie de mélange (« pooling ») des clones de la banque d'ADN génomique intégrés dans un vecteur selon l'invention. 8.A : mélange des clones de 24 plaques (« superpooling »). 8.B : définition de pools de lignes, colonnes et plaques individuelles. 8.C : plaque de PCR à partir d'un « superpool ». 8.D plaque
20 pour PCR à partir des pools de lignes, colonnes et plaques.

EXEMPLES

Exemple 1 :

- La première étape consiste à créer une banque de mini-BAC (entre 20 et 30
25 kb) adaptée à la construction de tous les types de vecteurs de recombinaison homologue. Cette banque est produite à partir d'ADN génomique de souris 129 sv et d'un vecteur pBeloBac11 modifié (figure 1). Des gènes de sélection négatifs sont rajoutés de part et d'autre du poly-linker de pBeloBAC. Les fragments d'ADN génomique dits grand bras et petit bras d'homologie du vecteur de recombinaison
30 homologue (Thomas et Capecchi, 1987 ; Hasty et al., 1991 ; Thomas et al., 1992) proviennent de cette banque.

Des techniques standard de sélection comme la PCR et le Southern blot sont utilisées pour sélectionner le mini-BAC contenant la partie à modifier du gène d'étude.

Des transposons sont insérés dans le mini-BAC afin d'introduire des gènes de sélection procaryotes, eucaryotes, et des sites d'action des recombinaisons. Ces transposons portent également des séquences spécifiques pouvant servir aux séquençages, aux tests PCR et aux digestions enzymatiques (Figure 2).

Le vecteur de recombinaison homologue est électroporé dans des cellules ES, l'événement de recombinaison homologue est sélectionné en présence de Zéocine, de G418 (Yagi et al., 1990) (Figure 3).

L'obtention de la mutation dite propre dans ces cellules ES recombinées est obtenue par l'action de la recombinaison FLP en présence de gancyclovir (Hasty and Bradley, 1993) (action d'élimination des cellules ES contenant le gène de la thymidine kinase TK).

Construction de la banque de BAC

Le génome de la souris est formé d'environ 3 milliards de paires de bases. Pour couvrir de façon satisfaisant ce génome, il est essentiel d'avoir une bonne représentation de sa diversité au sein de la banque de départ. Un modèle statistique a été élaboré dans ce but. Il se base sur l'équation mathématique suivante:

$$N = -L0 \ln ((1-p)/(L-l))$$

Cette équation prend en compte les éléments suivants :

- L : taille du fragment inséré dans le mini-BAC
- N : nombre de clones composant la banque
- p : pourcentage de réussite du criblage de la banque
- l : taille du locus cible devant être présent dans le mini-BAC
- L : Longueur minimum insécable d'intérêt
- L0 : longueur du génome

La taille du fragment génomique est choisie entre 20 et 30 kb (moyenne à 25 kb). Cette taille permet une manipulation facilitée des BAC tout en n'augmentant pas trop le nombre de clones devant composer la banque de BAC.

La première étape consiste à sélectionner le ou les mini-BACs positifs pour la région du gène d'intérêt (locus cible). Ce travail est réalisé par analyse de la totalité de la banque en associant des techniques simples de sélection par PCR et par Southern blot. En moyenne, 150 PCR sont nécessaires pour réaliser ce travail (test
5 des superpools puis des pools et enfin identification des clones positifs).

Parmi les BAC détectés positifs, il convient de choisir celui qui contient le gène d'intérêt dans une position adéquate. En effet, il est essentiel que les sites d'insertion visés pour l'intégration des transposons ne soit pas trop proche des extrémités. Le petit bras d'homologie doit avoir une taille minimale d'environ 3,5
10 Kb pour autoriser la construction du vecteur témoin positif et du vecteur final de recombinaison homologue. Ce choix se fait par simple profil enzymatique suivi d'une détection des fragments cibles.

Intégration des transposons

15 Dans le mini-BAC choisi, deux transposons (Tn5) sont intégrés aléatoirement, porteur chacun de gènes de résistances différents (Yang et al., 1997). Chacun possède des séquences spécifiques permettant de remplir sa mission d'insertion (OE de deux fois 19 Pb). Les insertions et la caractérisation du site d'insertion de chaque transposon se fait successivement.

20 *Le transposon 1* correspond à la cassette exogène classique d'un vecteur de recombinaison homologue (Figure 2) contenant le gène de sélection positive (Néo) et négatif (TK) encadré de sites d'action d'une recombinase ou non, de gènes marqueurs (lacZ par exemple) ou non ... Il est a noté que ce gène de sélection positif est également utilisé comme gène de sélection de l'intégration du transposon dans le
25 mini-BAC (promoteur mixte eucaryote-procaryote).

Pour l'inactivation classique d'un gène, le transposon 1 est introduit dans une région indispensable à l'expression du gène (exon 1 par exemple). Sinon il sera introduit dans un intron pour permettre une inactivation conditionnelle ultérieure.

30 *Le transposon 2* apporte un deuxième gène de sélection négative (TK) et un gène de sélection positive à la zéocine, encadré de sites d'action d'une recombinase. Ce transposon permet d'introduire une pression de sélection sur l'intégration d'un site loxP (additionné ou non d'autre séquence) utilisé par la recombinase Cre. De nombreux problèmes lors de l'événement de recombinaison homologue ont été

mentionnés pour les modèles animaux d'inactivation conditionnelle de gène. En effet, il n'est pas rare d'avoir une intégration homologue de la partie se trouvant sous la pression de sélection au G418 (néomycine) mais de perdre à chaque fois la séquence loxP (34 pb) présente en amont ou en aval. L'introduction du deuxième
5 gène de sélection (zéocine) au niveau de cette séquence loxP augmente les chances d'obtenir des cellules ES ayant intégré de façon homologue les deux transposons et donc les deux séquences loxP d'intérêt.

Les transposons sont conçus afin de pouvoir déterminer leur position relative ainsi que leur position par rapport aux extrémités de l'insert du mini-BAC
10 (site de restriction (S1, S2, S3, S4), séquences spécifiques pour amorces de PCR (OE)...).

Sélection de l'événement de recombinaison homologue

À ce stade, de nombreux antibiotiques sont utilisés en association.
15 L'intégration des deux séquences loxP se fait sous la pression de sélection au G418 et à la zéocine. La sélection de l'événement de recombinaison homologue est facilitée par la sélection négative DTA au deux extrémités de la construction.

20 "Nettoyage" des cellules ES recombinées : obtention d'une mutation dite "propre"

Pour produire un modèle murin le plus proche possible de la réalité, il est nécessaire d'enlever les gènes utilisés pour les sélections des événements d'intégration des transposons dans le mini-BAC et de l'événement d'intégration du transgène dans les cellules ES (Figure 4). Dans cette optique, il est intéressant
25 d'utiliser des séquences reconnues par la recombinase FLP, les sites FRT mutés ou non.

Le principe est le suivant : deux sites FRT vont permettre à la FLP d'exciser les séquences se trouvant entre ces deux sites. Il ne reste alors plus qu'un des deux sites FRT. Un site FRT muté ne peut agir qu'avec un autre site FRT muté (Schlake
30 and Bode, 1994). De cette manière, il nous est possible d'exciser les gènes de sélection de l'insertion des deux transposons. Il ne reste alors plus que les séquences permettant l'insertion des transposons (OE), les deux séquences loxP et deux séquences FRTs (mutées ou non). Ces séquences (OE, loxP, FRT) se retrouvent sur

deux sites d'insertion et représentent chacun 150pb (voir figure 4). Les cellules ES recombinaées dites propres sont micro-injectées dans des blastocystes pour donner un modèle murin.

5 Exemple 2 : Création du vecteur de clonage rTgV

• Construction du vecteur rTgV

- 1) Construction d'un dérivé de pBeloBAC11 ayant perdu le fragment compris entre les sites Sall 7030 et Sall 646 ainsi que ces sites Sall et contenant à la place du fragment compris entre les sites Sall 7030 et Sall 646 la séquence de clonage intermédiaire suivante (SEQ ID N° 2) :

5'ctcgagtaactataacgggtcctaaggtagcggcgcccatcgatgtcgactcgctaccttaggaccgtt
atagttactcgag – 3'

- contenant les sites de restriction suivant : XhoI - I-CeuI – AscI – ClaI – Sall-
I-CeuI- XhoI. Les extrémités XhoI ont été liguées avec les extrémités Sall du
vecteur.

Le plasmide résultant est appelé pBeloBAC13.

- 2) Construction d'un dérivé de pBeloBAC13 ayant perdu le site BstEII. Le vecteur pBeloBAC13 a été digéré par BstEII, traité à la Klenow et religué.

Le plasmide résultant est appelé pBeloBAC13 BK-1.

- 3) Construction d'un dérivé de pBeloBAC13 BK-1 contenant un site SwaI dans la séquence de clonage intermédiaire. Le vecteur pBeloBAC13 BK-1 a été digéré par AscI et Sall et la séquence suivante a été insérée entre ces deux sites :

5' – ggcgccgcatTTaaatctcgag – 3'

Cette séquence contient les sites de restriction suivant : AscI – SwaI – XhoI. L'extrémité XhoI a été liguée avec l'extrémité Sall du vecteur et l'extrémité AscI avec l'extrémité AscI du vecteur.

Le plasmide résultant est appelé pBeloBAC13 BK-1 AS-1.

- 4) Construction d'un dérivé du vecteur DTArTgV contenant un site SmaI à la place du site Sall en position 2 et un site Sall à la place du site BamHI en position 3049. Cette construction a été effectuée en insérant des oligonucléotides double brin

contenant les sites *Sma*I en position 2 et *Sal*I en position 3409 du plasmide DTArTgV.

Le plasmide résultant est appelé DTA Δ *Sal*I Δ *Bam*HI.

- 5 5) Clonage au site *Swa*I du vecteur pBeloBAC13 BK-1 AS-1 de deux copies du fragment de 3.4 kb contenant le gène DTA.

Une copie du gène DTA provient du vecteur DTA rTgV digéré par *Bam*HI, traité à la Klenow, digéré par *Sal*I et purification du fragment d'intérêt de 3.4 kb. L'autre copie du gène DTA provient du vecteur Δ TA Δ *Sal*I Δ *Bam*HI digéré
10 par *Sma*I, *Sal*I et purification du fragment d'intérêt de 3.4 kb. A la suite de ce clonage, l'une des extrémités *Swa*I du plasmide pBeloBAC13 BK-1 AS-1 est liguée à l'extrémité *Bam*HI (Klenow) d'une copie du gène DTA et l'autre extrémité *Swa*I est liguée à l'extrémité *Sma*I de la seconde copie du gène DTA. Les deux copies du gène DTA sont liées par leur extrémité *Sal*I.

- 15 Le plasmide résultant est appelé pBeloBAC13 BK-1 AS-1 DTA-2.

6) Préparation d'un plasmide dérivé de Bluescript KS(+) contenant la séquence suivante:

5'-

20 tcgagggccggccgagctcatgcattgcggccgcgtttaaacatttaaatacgactcactatagggcgaggatcc
aagcttagtattctatagtgacacctaataatcgatgtcgaccggaccggcccgggcgcatgcttaattaatggcaaacagc
tattatgggtattatgggtctcgag-3'

Cette séquence contient les sites de restriction suivant dans cet ordre : *Xho*I
– *Fse*I – *Sac*I – *Nsi*I – *Not*I – *Pme*I – *Swa*I – *Bam*HI – *Hind*III – *Sal*I – *Rsr*II – *Srf*I –
25 *Sph*I – *Pac*I – *Pi*-*Psp*I – *Xho*I ainsi qu'une séquence dite T7 entre les sites de
restriction *Swa*I et *Bam*HI et une séquence dite SP6 entre les sites de restriction
*Hind*III et *Sal*I. Cette séquence est désignée par [polylinker] dans la suite de la
description.

Le plasmide résultant est appelé BlueLC-2.

30

7) Clonage au site *Sal*I du plasmide pBeloBAC13 BK-1 AS-1 DTA-2 du «
polylinker » préparé par digestion du plasmide BlueLC-2 par *Xho*I et purification
du fragment d'intérêt de 187 pb.

Le plasmide résultant est appelé pBeloBAC13 BK-1 AS-1 DTA-2 LC-16.

Exemple 3 : Création de la banque de BAC utilisant le vecteur de clonage rTgV décrit dans l'exemple 2.

5 Cette banque est constituée de 300 000 clones représentant 2,5 fois le génome murin. Elle a été réalisée selon les procédures suivantes :

- **Souche bactérienne et conditions de culture**

E. Coli DH10B (Grant et al. 1990) a été utilisée comme hôte des mini-BAC.

10 Le vecteur de base (vecteur genOway décrit plus haut) a été utilisé pour construire cette banque. La culture des bactéries DH10B s'est effectué de manière classique dans du LB à 37°C (Sambrook et al. 1989). Les clones recombinants ont été sélectionnés sur des boîtes d'agar contenant 12,5 µg/ml de chloramphenicol à partir de E. Colis DH10B électrocompétentes non congelées (Sheng et al. 1995).

15

- **Préparation d'ADN génomique et construction de la banque**

L'ADN génomique a été préparé à partir de culture de cellules souches embryonnaires (ES) de souris en fond génétique 129svJ selon le protocole standard décrit dans Sambrook et al. (Chapitre 9) et dans Vaiman et al. (1999).

20 L'ADN génomique est digéré partiellement avec HindIII, les fragments de 20-30 kb sont isolés par Pulse field et introduit dans le vecteur rTgV par ligation HindIII.

La préparation du vecteur rTgV a été réalisée selon le protocole de Woo et al., 1994.

25

Exemple 4 : Elaboration de la structure de la banque de BAC et de son mode de criblage

Après obtention des différents clones composant la banque de BAC, contenant l'ADN génomique, la banque est classée, afin de pouvoir être rapidement
30 criblée pour identifier le(s) clone(s) portant le locus d'intérêt auquel on désire effectuer l'opération de recombinaison.

On repique donc les colonies et l'on effectue différentes étapes de mélanges entre les colonies (« pooling »), ce qui permet d'effectuer moins de réaction de criblage et de pouvoir malgré tout identifier les clones de façon univoque.

5 • Repiquage des colonies

Les colonies préalablement étalées sur plateau 22x22 cm² sont repiquées par une station robotisée 'Qpix' de la marque Genetix puis déposées dans des plaques 96 puits (marque TPP distribuées par ATGC réf. T92697) contenant 200 µl de milieu 2YT (Sigma réf. Y2377) - glycerol 10% final - chloramphénicol 12.5 µg/ml final
10 (Sigma réf. C0378) par puits (plaques MBAA 0001 à 3168 'A'). Ces plaques sont incubées toute la nuit à 37°C

• Copies des plaques de la banque

Deux copies de ces plaques sont réalisées :

- 15 - Une copie en plaque de 96 puits identique à la plaque mère
 - Une copie en plaques de 96 puits de type « deep-well » (ATGC réf. 219009) contenant 1 ml de milieu 2YT-chloramphénicol par puits. Cette plaque servira à la manipulation de pooling des clones. Ces plaques sont incubées toute la nuit à 37°C.

20

• Pooling des clones

Les plaques « deep-well » sont groupées par 24.

Le premier niveau de pooling est la réalisation de « Superpools ». Ils correspondent au mélange de tous les clones d'un groupe de 24 plaques. Le nombre
25 de superpools est donné par le nombre total de plaques de la banque divisé par 24. Ce système permet de localiser rapidement un clone dans un groupe de 24 plaques en effectuant un nombre de points de PCR correspondant au nombre de superpools (figure 8 .A).

Le deuxième niveau de pooling est la réalisation de 'pools'. Toujours par
30 groupes de 24 plaques, les clones sont mélangés par lignes (pools de ligne) par colonnes (pools de colonnes) et par plaques (pools de plaques), ainsi qu'indiqué sur la figure 8.B. Si un clone se trouve dans ce groupe, un des 24 pools de plaque doit

être positif en PCR ainsi qu'un pool de ligne et de colonne. Le recoupement de ces trois informations donne l'adresse du clone dans la banque.

Ces mélanges sont réalisés à l'aide d'une station robotisée 'Biorobot3000' de la marque Qiagen ou une station robotisée 'Biomek2000' de la marque Beckman.

- 5 Ces stations sont des robots 'pipeteurs' fonctionnant à l'aide de cônes stériles jetables.

- Préparation des ADN des pools

- 10 Tous les mélanges sont centrifugés, les culots bactériens lavés avec du Tris-EDTA (sigma réf. E5134) (TE), recentrifugés puis resuspendus dans du TE. L'ADN des BACs est ensuite récupéré en faisant bouillir les mélanges au micro-ondes. Puis les mélanges sont centrifugés et on récupère les surnageants.

- Distribution des ADN en plaque 96 puits pour le criblage par PCR

- 15 Les surnageants de tous les superpools sont distribués dans une ou plusieurs plaques de 96 puits deep-well afin de constituer un stock qui sera facilement manipulable à l'aide d'une pipette multicanaux (Figure 8.C).

- 20 Pour un pool, les 24 surnageants des mélanges de plaque, les 12 surnageants des mélanges de colonnes et les 8 surnageants des mélanges de ligne sont distribués dans une demi plaque de 96 puits deep-well (Figure 8.D). Dans chaque plaque on peut disposer deux pools.

- 25 L'ADN de chaque puits peut être ensuite dilué pour préparer des plaques utilisables en PCR.

Exemple 5 : Réalisation d'un vecteur de ciblage et recombinaison homologue en cellule ES

- Vérification de la stabilité des mini-BAC de la banque

- 30 Un échantillon représentatif de 40 clones a été cultivé et analysé de manière à vérifier la stabilité des mini-BAC. Aucune instabilité ou diminution de la croissance des bactéries n'a été observée par rapport au vecteur contrôle pBeloBac11 sans insert.

- Vérification de la taille des inserts de la banque de BAC

Les ADN plasmidiques de 40 clones ont été analysés par digestion enzymatique. Les tailles des inserts se trouvent toutes entre 20 et 30 Kb (comme
5 décrit dans les conditions optimum de la création de la banque rTgV).

- Construction du vecteur de ciblage

Dans ces 40 clones, un mini-Bac contenant un insert de 20 kb a été choisi. Un gène de sélection a été introduit par digestion enzymatique et religation. Le petit
10 d'homologie (1,4 Kb pour le témoin positif, 1,1Kb pour le vecteur de RH) et le grand bras d'homologie (9Kb) ont été obtenus par action de l'exonucléase III (voir protocole dans Sambrook et al, Chapitre 5). Un vecteur dit témoin positif et un vecteur de recombinaison homologue ont ainsi été construits. Un simple séquençage (de deux fois 500 pb) a suffi pour élaborer le criblage de l'événement de
15 recombinaison homologue.

- Electroporation et recombinaison homologue en cellule ES

Cette étape est réalisée selon des méthodes connues dans l'art. Le taux de recombinaison homologue en cellule ES dépend des séquences d'homologie du
20 vecteur de recombinaison homologue et non de son mode d'élaboration.

REFERENCES

- Araki, *et al.* (1995). Proc Natl Acad Sci U S A 92, 160-4.
Austin, *et al.* (1981). Cell 25, 729-36.
25 Babinet, et Cohen-Tannoudji,. (2000). médecine/science 16, 31-42.
Baron U. et al, (1999). PNAS 96 : 1008-1013.
Brune, *et al.* (1999). Nat Biotechnol 17, 360-4.
Capecchi,. (1989). Science 244, 1288-92.
Chatterjee, et Coren, (1997).. Nucleic Acids Res 25, 2205-12.
30 Cohen-Tannoudji et Babinet, (1998). Mol Hum Reprod 4, 929-38.
Cohen-Tannoudji, *et al.* (1998). Mol Cell Biol 18, 1444-8.
Colleaux, *et al.* (1988). Proc Natl Acad Sci U S A 85, 6022-6.
de Wit, *et al.* (1998). Nucleic Acids Res 26, 676-8.

- Detloff, et al. (1994). *Mol Cell Biol* 14, 6936-43.
- Evans, et Kaufman,. (1981). *Nature* 292, 154-6.
- Fiering,et al. (1999). *Methods Enzymol* 306, 42-66.
- Goryshin, et al. (2000). *Nat Biotechnol* 18, 97-100.
- 5 Goryshin et Reznikoff, (1998). *J Biol Chem* 273, 7367-74.
- Grant et al 1990. *Proc. Natl. Acad. Sci. USA* 87, 4645-4649.
- Gu, et al. (1994). *Science* 265, 103-6.
- Harada, et al. (1999). *Embo J* 18, 5931-42.
- Hasty, et Bradley. (1993). Gene targeting vectors for mammalian. In *Gene*
- 10 *targeting: A pratical Approach*, I. press, ed., pp. 1-31.
- Hasty et al. (1991). *Nature* 350, 243-6.
- Hasty, et al. (1991). *Mol Cell Biol* 11, 5586-91.
- Hoess et Abremski (1985). *J Mol Biol* 181, 351-62.
- Jaxson laboratory. The transgenic/ Targeted Mutated Data base.
- 15 Kilby, et al. (1993). *Trends Genet* 9, 413-21.
- Kulkarni, et al. (1999). *Cell* 96, 329-39.
- laboratory, C. protocols/BAC library/sequences.
- laboratory, C. protocols/pictures/pbeloBAC.
- laboratory, C. G. R. protocols/BAC library/construction.
- 20 Li, et al. (1997). *J Cell Biol* 139, 129-44.
- Marshall, et Lemieux. (1991). *Gene* 104, 241-5.
- Martin, (1981). *Proc Natl Acad Sci U S A* 78, 7634-8.
- Monteilhet, et al. (1990). *Nucleic Acids Res* 18, 1407-13.
- Moore, et al. (1998). *Nat Genet* 18, 118-25.
- 25 Muyrers, et al. (1999). *Nucleic Acids Res* 27, 1555-7.
- Sambrook, et al , 1989. *Molecular Cloning: A Laboratory Manual*, second ed. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.
- Sauer, et Henderson, (1988). *Proc Natl Acad Sci U S A* 85, 5166-70.
- Schibler et al. (1998). *Mamm Genome* 9, 119-24.
- 30 Schlake, et Bode, (1994). *Biochemistry* 33, 12746-51.
- Schneider-Maunoury, et al. (1993). *Cell* 75, 1199-214.
- Sheng et al, 1995. *Nucleic Acids Res.* 23, 1990-1996.
- Shibata, et al. (1997). *Science* 278, 120-3.

- Smithies, (1993). Animal models of human genetic diseases. *Trends Genet* 9, 112-6.
- Smithies, et al. (1985). *Nature* 317, 230-4.
- Stellwagen, et Craig, (1997). *Embo J* 16, 6823-34.
- Sternberg, et al. (1986). *J Mol Biol* 187, 197-212.
- 5 Sunaga, et al. (1997). *Mol Reprod Dev* 46, 109-13.
- Tajbakhsh, et al. (1996). *Nature* 384, 266-70.
- Thomas, et Capecchi, (1987). *Cell* 51, 503-12.
- Thomas, et al. (1992). *Mol Cell Biol* 12, 2919-23.
- Tsien et al. (1996).. *Cell* 87, 1327-38.
- 10 Vaiman D ; et al. (1999). *Mammalian Genome* 10, 585-587
- Westphal et Leder. (1997). *Curr Biol* 7, 530-3.
- Wong, et al. (1986). *Somat Cell Mol Genet* 12, 63-72.
- Woo, S. et al. (1994) *Nucleic Acids Res.* 22, 4922-4931
- Xu et al. (1999). *Nat Genet* 22, 37-43.
- 15 Yagi, et al. (1990). *Proc Natl Acad Sci U S A* 87, 9918-22.
- Yang, et al. (1997).. *Nat Biotechnol* 15, 859-65.

Revendications

1. Vecteur de clonage adapté pour la création d'une banque d'ADN génomique, caractérisé en ce qu'il présente une cassette comprenant
 - 5 - un polylinker permettant l'intégration d'un fragment d'ADN génomique, ledit polylinker étant flanqué de sites rares de restriction et que ladite cassette est flanquée de sites très rares de restrictions.
2. Vecteur selon la revendication 1, caractérisé en ce que ladite cassette comprend également au moins un gène de sélection négative.
- 10 3. Vecteur selon la revendication 1 ou 2, caractérisé en ce ledit polylinker permettant l'intégration d'un fragment d'ADN génomique présente un maximum de 4 sites de restrictions.
4. Vecteur selon l'une des revendications 1 à 3, caractérisé en ce qu'il est dérivé d'un cosmide, ou d'un vecteur de chromosome artificiel, de préférence de
 - 15 bactérie.
5. Vecteur selon la revendication 4, caractérisé en ce qu'il est dérivé du vecteur pBeloBAC 11.
6. Vecteur selon la revendication 4 ou 5, caractérisé en ce qu'au moins un site de restriction choisi parmi les sites présents 1, 2 ou 3 fois dans le squelette du
 - 20 vecteur parent en est absent.
7. Vecteur selon l'une des revendications 5 ou 6, caractérisé en ce que son squelette ne présente plus au moins un site choisi parmi *ApaI*, *BstEII*, *SacII*, *SfiI*, *SpeI*, *SphI*, *StuI*, *XhoI*, *BssHII*, *EcoRI*, *EcoRV*, *KpnI*, *NdeI*, *NotI*, *NruI*, *PvuI*, *SgrAI*, *XbaI*, *PstI*, *SalI* et *SmaI*.
- 25 8. Vecteur selon la revendication 7, caractérisé en ce que son squelette ne présente plus de site *BstEII*.
9. Vecteur selon l'une des revendications 1 à 8, caractérisé en ce que ladite cassette présente un site très rare de restriction à l'une des ses extrémités et deux sites très rares de restriction à l'autre extrémité, l'un de ces deux sites étant
 - 30 identique au site situé à l'autre extrémité de la cassette.
10. Vecteur selon l'une des revendications 1 à 9, caractérisé en ce que ledit polylinker est flanqué de plusieurs (nombre supérieur à 3) sites rares de

restrictions choisis notamment parmi *PmeI*, *SgrAI*, *RsrII*, *ClaI*, *NofI*, *PacI*, *SrfI*, *NheI*.

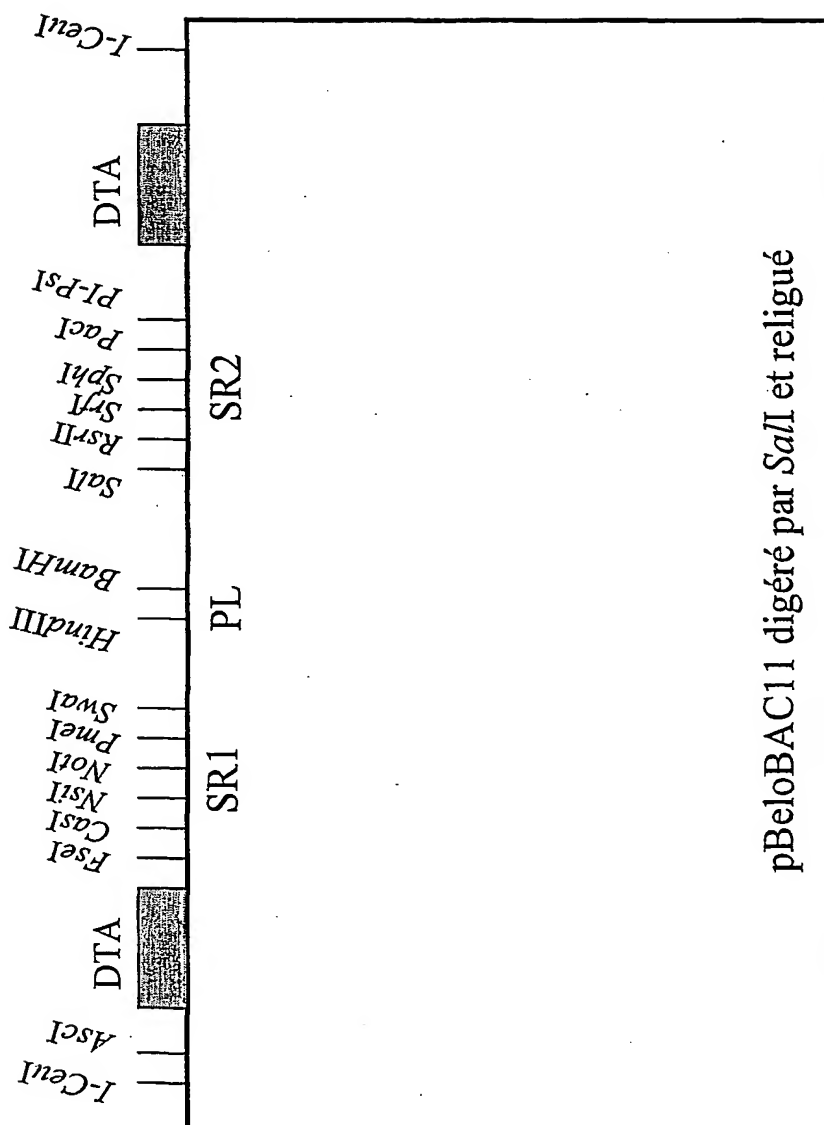
11. Vecteur de clonage selon l'une des revendications 2 à 10, caractérisé en ce que ladite cassette comprend deux copies du même gène de sélection négative,
5 flanquant ledit polylinker et sites rares de restriction.
12. Vecteur selon l'une des revendications 1 à 11, caractérisé en ce qu'il s'agit du vecteur de séquence SEQ ID N° 1.
13. Procédé pour effectuer une recombinaison ciblée dans un organisme, à partir d'une banque d'ADN génomique dans un vecteur selon l'une des revendications
10 1 à 12 comprenant :
 - a) l'incubation d'un vecteur de ladite banque génomique avec au moins un fragment d'ADN comprenant au moins un transposon, ledit transposon comprenant
 - un gène de sélection positive,
 - 15 - éventuellement un gène de sélection négative et/ou un gène marqueur,lesdits gènes de sélection étant éventuellement encadrés par les sites d'action I d'une recombinaise site-spécifique I,
 - et de façon optionnelle au moins un site d'action II d'une
20 recombinaise site-spécifique II différente de la première recombinaise I,
ledit site d'action II n'étant pas situé entre lesdits sites d'action I de la première recombinaise I,
en présence d'une enzyme ayant une activité de transposase pour
25 le(s)dit(s) transposon(s), de telle sorte que le(s)dit(s) transposon(s) est (sont) transféré(s) dans ledit fragment génomique,
 - b) l'introduction dudit vecteur, éventuellement linéarisé, portant ledit fragment génomique dans lequel est (sont) inséré(s) le(s)dit(s) transposons dans une cellule hôte cible issue dudit organisme,
 - 30 c) la sélection des événements de recombinaison homologue dans ladite cellule cible par l'utilisation du (des) gène(s) de sélection positive porté(s) par le(s)dit(s) transposon(s), et éventuellement du gène de sélection négative porté par ledit vecteur de clonage.

14. Procédé selon la revendication 13, caractérisé en ce que l'étape c) est suivie d'une étape d) :
- 5 d) la soumission des cellules transformées par recombinaison homologue et sélectionnées dans l'étape c) à l'action de la recombinaise I, afin d'éliminer le(s) gène(s) de sélection positive et éventuellement négative porté(s) par le(s) transposon(s).
15. Procédé selon la revendication 13 ou 14, caractérisé en ce que l'organisme cible est un organisme pluricellulaire, et que ladite cellule cible issue dudit organisme est une cellule souche.
- 10 16. Procédé selon la revendication 15, caractérisé en ce que ledit organisme appartient au genre des rongeurs, notamment à l'espèce murine.
17. Procédé selon l'une des revendications 13 à 16, caractérisé en ce que l'on induit la délétion d'un fragment génomique dans ledit organisme après l'étape c) ou éventuellement l'étape d) par la mise en contact des cellules sélectionnées dans
- 15 l'étape c) ou dérivées desdites cellules avec la recombinaise II.
18. Kit comprenant :
- un vecteur de clonage selon l'une des revendications 1 à 12,
 - au moins un fragment d'ADN comprenant au moins un transposon, ledit transposon comprenant un gène de sélection positive, éventuellement un

20 gène de sélection négative et/ou un gène marqueur, éventuellement encadrés par les sites d'action I d'une recombinaise site-spécifique I, optionnellement au moins un site d'action II d'une recombinaise site-spécifique II différente de la première recombinaise I, ledit site d'action II n'étant pas situé entre lesdits sites d'action I de la première

25 recombinaise I.
19. Kit selon la revendication 18, caractérisé en ce qu'il comprenant un ou deux fragments d'ADN, chacun comprenant un transposon, chaque transposon portant un gène de sélection positive différent, et éventuellement un gène de sélection négative, de préférence identique pour les deux transposons, lesdits
- 30 gènes de sélection positive et négative étant encadrés par les sites d'action d'une recombinaise I.

20. Kit selon la revendication 19, caractérisé en ce que chaque transposon comprend une séquence correspondant à un site d'action d'une recombinaise II différente de la première recombinaise I.
21. Kit selon l'une des revendications 18 à 20, caractérisé en ce que ladite
5 recombinaise I est la recombinaise FLP, ladite recombinaise II étant la recombinaise Cre.
22. Kit selon l'une des revendications 18 à 21, caractérisé en ce qu'il comprend une enzyme ayant une activité de transposase pour le(s)dit(s) transposon(s).
23. Kit selon l'une des revendications 18 à 22 pour la mise en œuvre d'un procédé
10 selon l'une des revendications 13 à 17.
24. Banque d'ADN génomique, caractérisée en ce que les fragments d'ADN génomique sont clonés dans un vecteur selon l'une des revendications 1 à 12.



1
 11

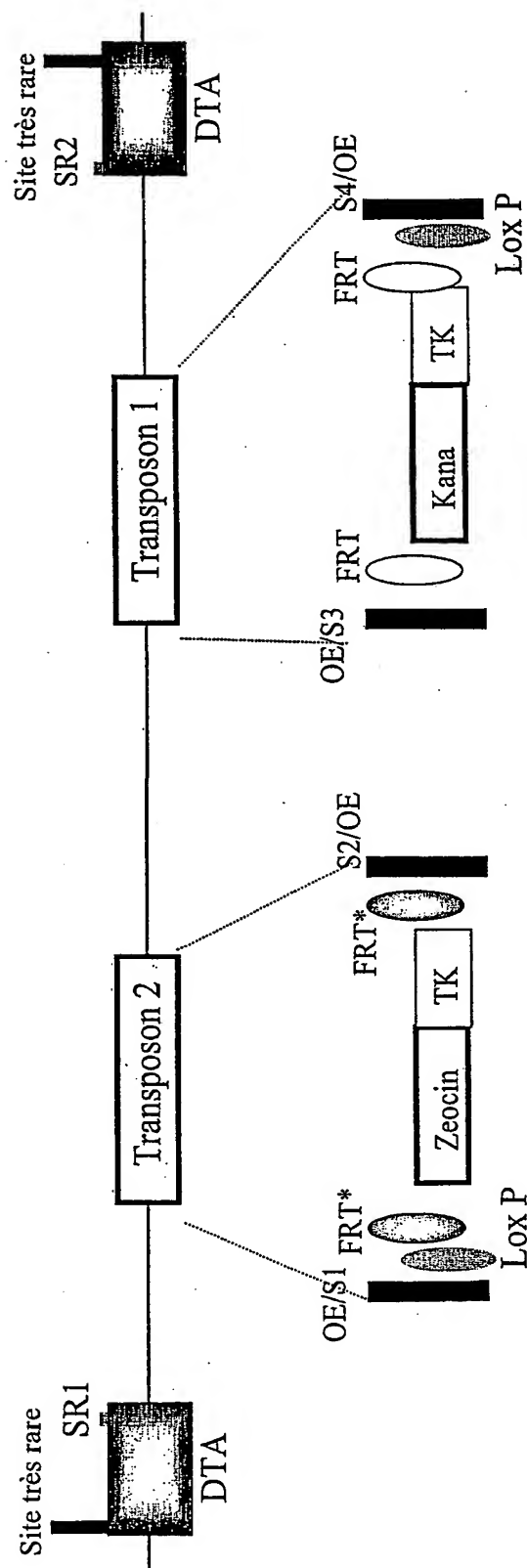


Fig. 2

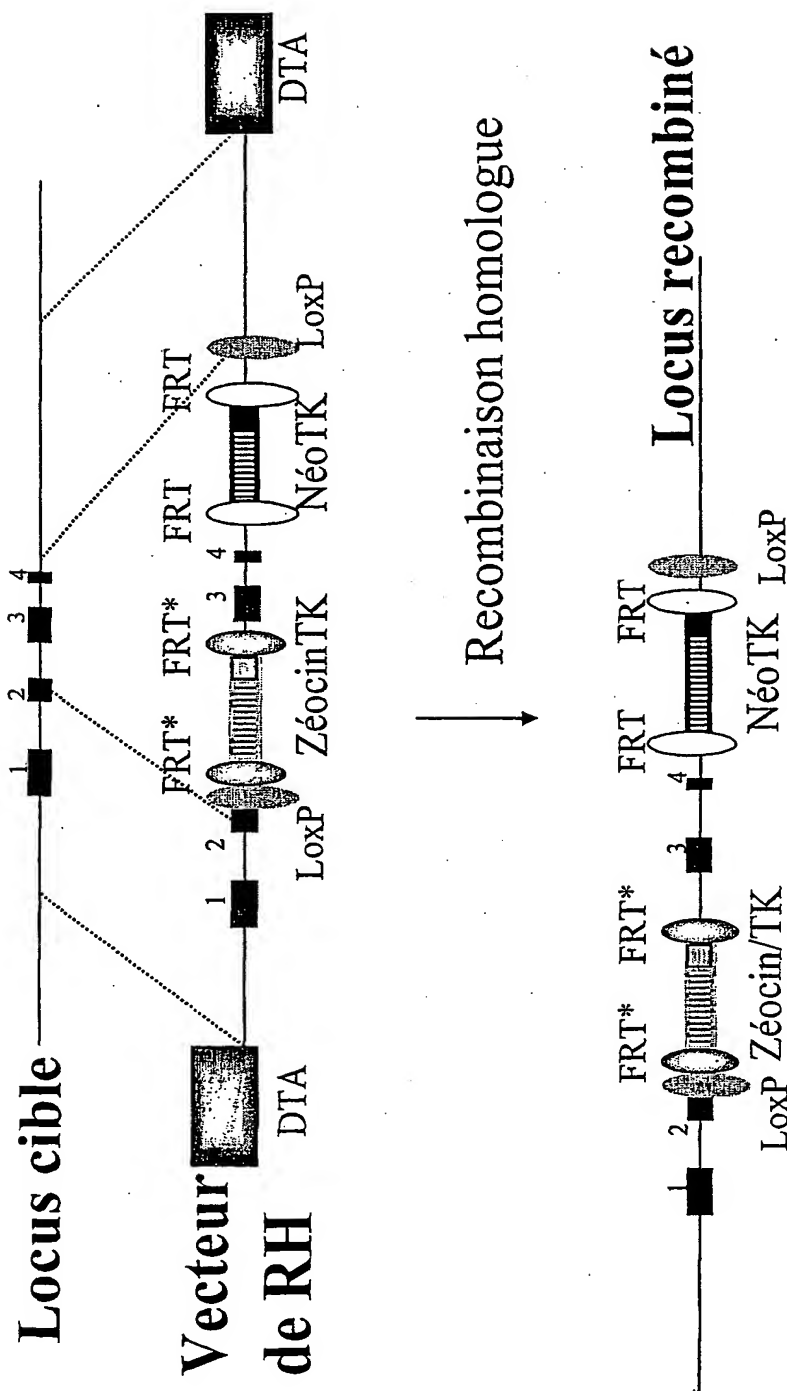


Fig. 3

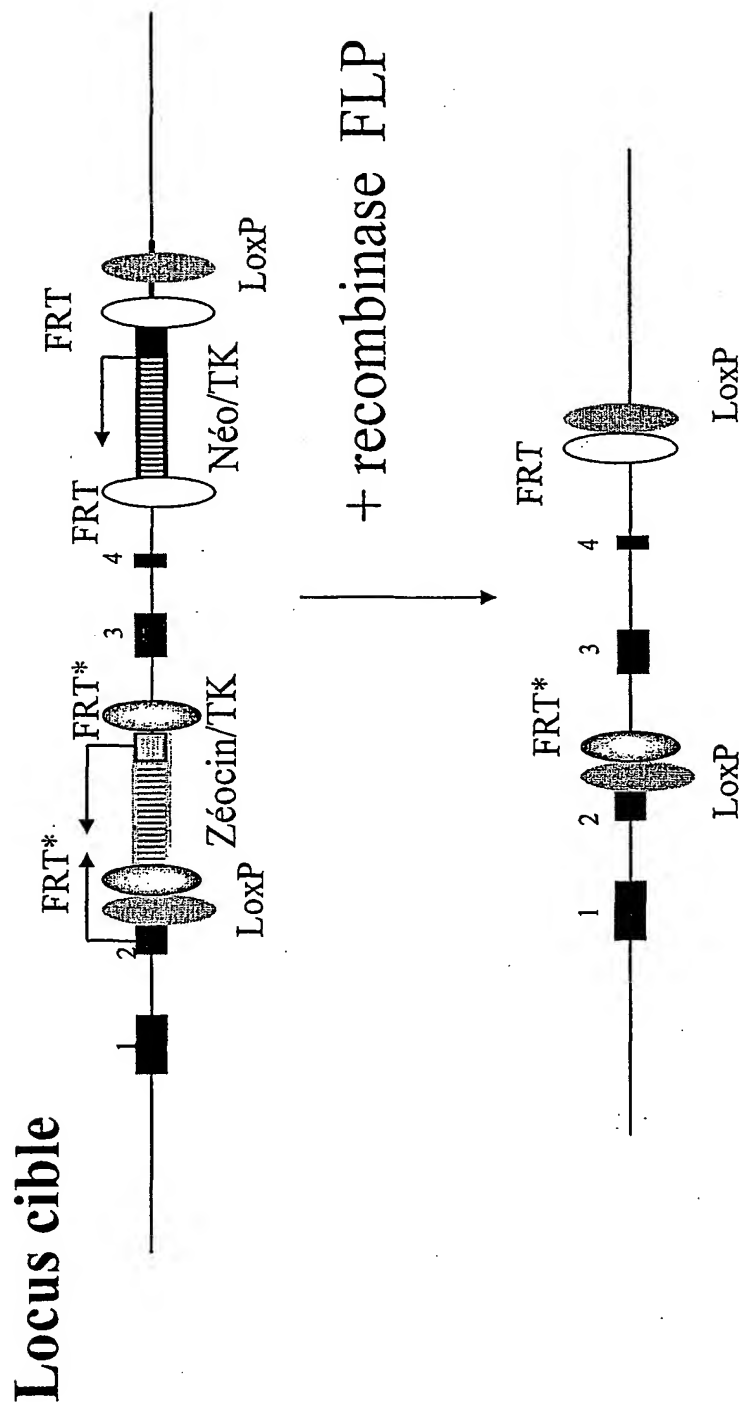


Fig. 4

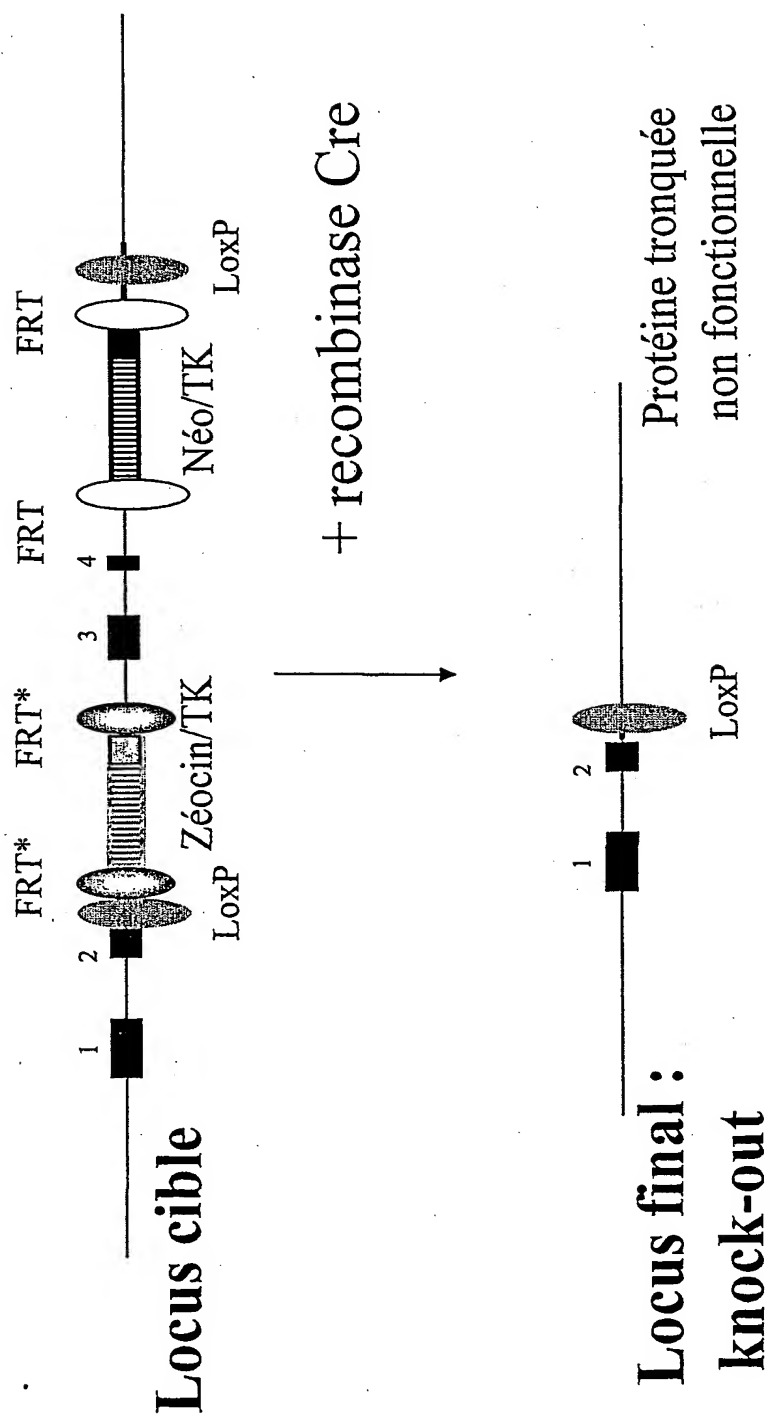


Fig. 5

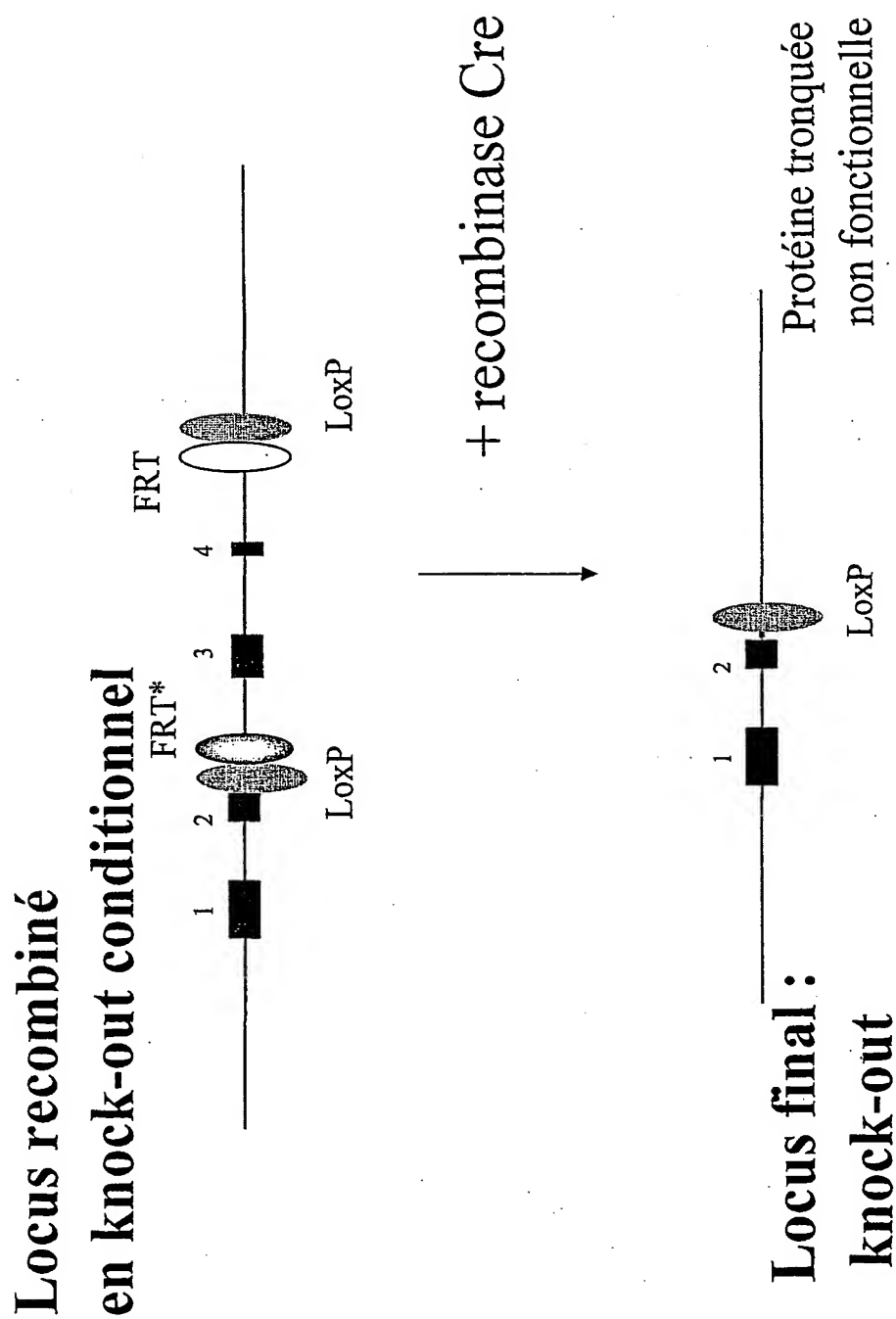


Fig. 6

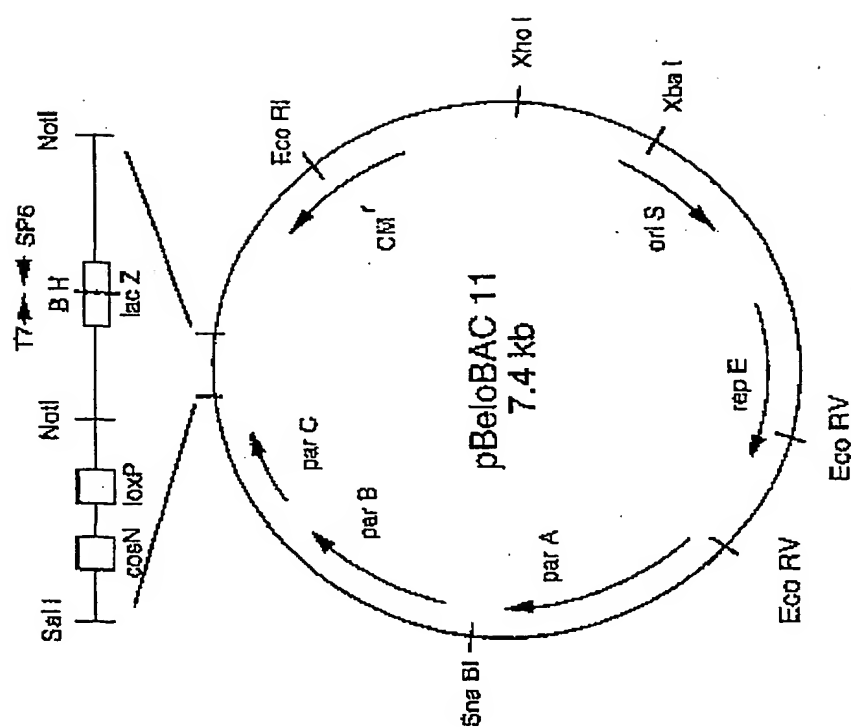


Fig. 7

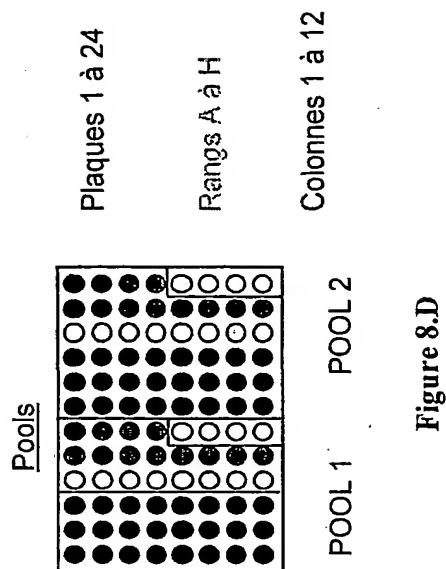
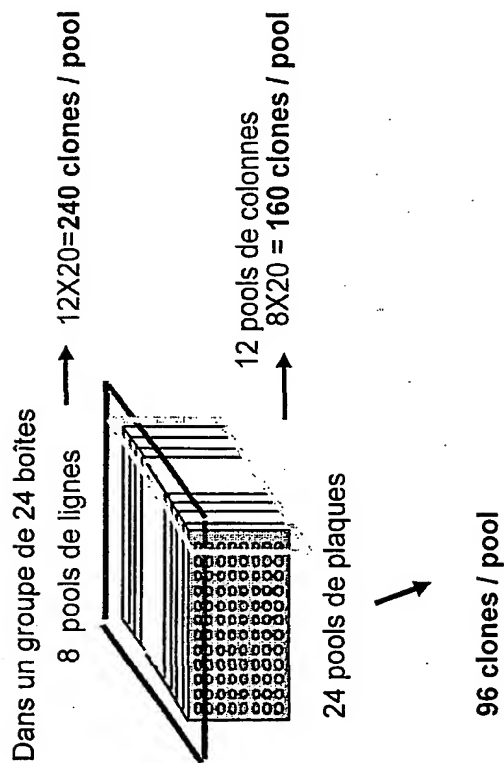
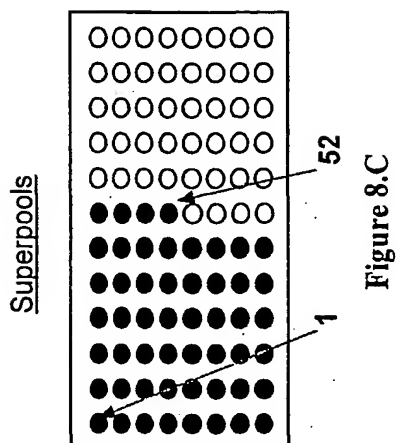
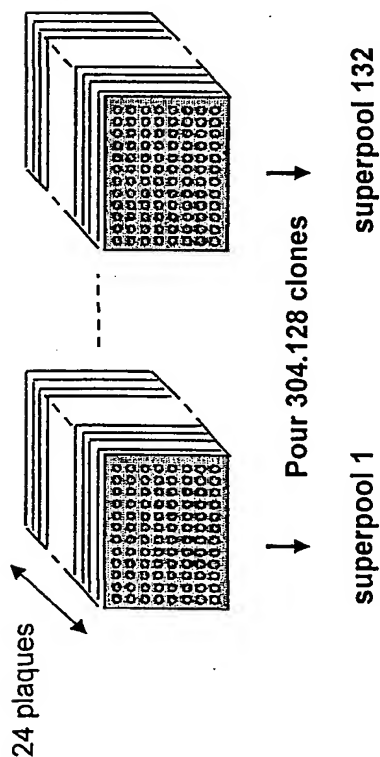


Fig. 8

LISTE DE SEQUENCES

<110> GENOWAY

<120> VECTEURS DE CLONAGE POUR RECOMBINAISON HOMOLOGUE ET
PROCEDE LES UTILISANT

<130> D19317

<150> FR 01/06921

<151> 2001-05-28

<160> 4

<170> PatentIn Ver. 2.1

<210> 1

<211> 13478

<212> ADN

<213> Séquence artificielle

<220>

<223> Description de la séquence artificielle: Vecteur
rTgV, 13478 bp, circulaire

<400> 1

```

ctcgaccaat tctcatgttt gacagcttat catcgaattt ctgccattca tccgcttatt 60
atcacttatt caggcgtagc aaccaggcgt ttaagggcac caataactgc cttaaaaaaa 120
ttacgccccg ccctgccact catcgcagta ctgttgtaat tcattaagca ttctgcccgc 180
atggaagcca tcacaaacgg catgatgaac ctgaatcgcc agcggcatca gcaccttgtc 240
gccttgcgta taatattttgc ccatggtgaa aacgggggcg aagaagttgt ccatattggc 300
cacgtttaaa tcaaaactgg tgaaactcac ccagggtattg gctgagacga aaaacatatt 360
ctcaataaac cctttaggga aataggccag gttttcaccc taacacgcca catcttgcca 420
atatatgtgt agaaactgcc ggaaatcgct gtggtattca ctccagagcg atgaaaacgt 480
ttcagttttg tcctggaaaa cgggtgaaca aggggtgaaca ctatcccata tcaccagctc 540
accgtctttc attgccatac ggaattccgg atgagcattc atcaggcggg caagaatgtg 600
aataaaggcc ggataaaaact tgtgcttatt tttctttacg gtctttaaaa aggccgtaat 660
atccagctga acggtctggt tataggtaca ttgagcaact gactgaaatg cctcaaaatg 720
ttctttacga tgccattggg atatatcaac ggtggtatat ccagtgatct ttttctccat 780
tttagcttcc ttagctcctg aaaatctcga taactcaaaa aatacgcccg gtagtgatct 840
tatttcatta tggtgaaagt tggaaacctc tacgtgccga tcaacgtctc attttcgcca 900
aaagttggcc cagggtctcc cgggtatcac agggacacca ggatttattt attctgcgaa 960
gtgatcttcc gtcacaggtt tttattcgcg aagtgcacga cagaacggtc aggacctgga 1080
caggaaggac agagaaagcg cggatctggg aagtgcacga cagaacggtc aggacctgga 1080
ttggggaggc ggttgccgcc gctgctgctg acggtgtgac gttctctgtt ccggtcacac 1140
cacatacgtt ccgccattcc tatgcatgac acatgctgta tgccggtata ccgctgaaag 1200
ttctgcaaag cctgatggga cataagtcca tcagttcaac ggaagtctac acgaagggtt 1260
ttgcgctgga tgtggctgcc cggcaccggg tgcagtttgc gatgccggag tctgatgcgg 1320
ttgcgatgct gaaacaatta tcctgagaat aaatgccttg gcctttatat ggaaatgtgg 1380
aactgagtgg atatgctggt tttgtctggt aaacagagaa gctggctgtt atccactgag 1440
aagcgaacga aacagtcggg aaaatctccc attatcgtag agatccgcat tattaatctc 1500
aggagcctgt gtagcgttta taggaagtag tgttctgtca tgatgcctgc aagcggtaac 1560
gaaaacgatt tgaatatgcc ttcaggaaca atagaaatct tcgtgcggtg ttacgttgaa 1620
gtggagcggg ttatgtcagc aatggacaga acaacctaat gaacacagaa ccatgatgtg 1680
gtctgtcctt ttacagccag tagtgctcgc cgcagtcgag cgacagggcg aagccctcga 1740
gtgagcaggg aagcaccagg gaacagcact tatatatctt gcttacacac gatgcctgaa 1800
aaaacttccc ttgggggttat ccacttatcc acggggatat ttttataatt atttttttta 1860
tagtttttag atcttctttt ttagagcgcc ttgtaggcct ttatccatgc tggttctaga 1920
gaaggtgttg tgacaaattg ccctttcagt gtgacaaatc accctcaaat gacagtcctg 1980
tctgtgacaa attgccctta accctgtgac aaattgcctt cagaagaagc tgttttttca 2040
caaagttatc cctgcttatt gactcttttt tatttagtgt gacaatctaa aaacttgtca 2100

```


cacttcacat	ggatctgtca	tggcggaaac	agcggttatc	aatcacaaga	aacgtaaaaa	2160
tagccccgga	atcgtccagt	caaacgacct	cactgaggcg	gcatatagtc	tctccccgga	2220
tcaaaaacgt	atgctgtatc	tgttcgttga	ccagatcaga	aaatctgatg	gcaccctaca	2280
ggaacatgac	ggatatctgcg	agatccatgt	tgctaaatat	gctgaaatat	tcggattgac	2340
ctctgcggaa	gccagtaagg	atatacggca	ggcattgaag	agtttcgcgg	ggaaggaagt	2400
ggttttttat	cgccctgaag	aggatgccgg	cgatgaaaaa	ggctatgaat	cttttccttg	2460
gtttatcaaa	cgtgcgcaca	gtccatccag	agggctttac	agtgtacata	tcaaccata	2520
tctcattccc	ttctttatcg	ggttacagaa	ccggttttacg	cagtttcggc	ttagtgaac	2580
aaaagaaatc	accaatccgt	atgccatgcg	tttatacgaa	tccctgtgtc	agtatcgtaa	2640
gcgggatggc	tcaggcatcg	tctctctgaa	aatcgactgg	atcatagagc	gttaccagct	2700
gcctcaaagt	taccagcgta	tgccctgactt	ccgccgcgcg	ttcctgcagg	tctgtgttaa	2760
tgagatcaac	agcagaactc	caatgcgcct	ctcatacatt	gagaaaaaga	aaggccgcca	2820
gacgactcat	atcgtatttt	ccttccgcga	tatcacttcc	atgacgacag	gatagtctga	2880
gggttatctg	tcacagattt	gaggggtggtt	cgctcacattt	gttctgacct	actgagggtg	2940
atltgtcaca	gttttgctgt	ttccttcagc	ctgcatggat	tttctcatac	tttttgaact	3000
gtaattttta	aggaagccaa	atltgagggc	agtttgctcac	agttgatttc	cttctctttc	3060
ccttcgtcat	gtgacctgat	atcgggggtt	agttcgtcat	cattgatgag	ggttgattat	3120
cacagtttat	tactctgaat	tggctatccg	cggtgtgtacc	tctacctgga	gtttttccca	3180
cgggtgatat	ttcttcttgc	gctgagcgta	agagctatct	gacagaacag	ttcttctttg	3240
cttctcgcgc	agttcgctcg	ctatgctcgg	ttacacggct	gcggcgagcg	ctagtataa	3300
taagtgcagc	aggtatgtgc	tcttcttate	tctttttgta	gtgttgctct	tattttaaac	3360
aactttgcgg	ttttttgatg	actttgcgat	tttgtgtgtg	ctttgcagta	aattgcaaga	3420
tttaataaaa	aaacgcaaag	caatgattaa	aggatgttca	gaatgaaact	catggaaaca	3480
cttaaccagt	gcataaacgc	tggtcatgaa	atgacgaagg	ctatcgccat	tgacagttt	3540
aatgatgaca	gcccggaagc	gaggaataa	acccggcgct	ggagaatagg	tgaagcagcg	3600
gatttagttg	gggtttcttc	tcaggctatc	agagatgccg	agaaagcagg	gcgactaccg	3660
caccgcgata	tggaaattcg	aggacgggtt	gagcaacgtg	ttggttatac	aattgaacaa	3720
attaatcata	tgcgtgatgt	gtttggtacg	cgattgcgac	gtgctgaaga	cgtatttcca	3780
ccggtgatcg	gggttgctgc	ccataaagg	ggcggtttaca	aaacctcagt	ttctgttctc	3840
cttgctcagg	atctggctct	gaaggggcta	cgtgttttgc	tcgtggaagg	taacgacccc	3900
cagggaacag	cctcaatgta	tcacggatgg	gtaccagatc	ttcataattca	tgacagaagc	3960
acttctctgc	ctttctatct	tggggaaaag	gacgatgtca	cttatgcaat	aaagccact	4020
tgctggccgg	ggcttgacat	tattccttcc	tgtctggctc	tgacaccgat	tgaactgag	4080
ttaatgggca	aatttgatga	aggtaaactg	cccaccgatc	cacacctgat	gctccgactg	4140
gccattgaaa	ctgttgctca	tgactatgat	gtcatagtta	ttgacagcgc	gcctaacctg	4200
ggtatcggca	cgattaatgt	cgtatgtgct	gctgatgtgc	tgattgttcc	cacgcctgct	4260
gagttgtttg	actacacctc	cgcactgcag	tttttcgata	tgcttcgtga	tctgctcaag	4320
aacgttgatc	ttaaagggtt	cgagcctgat	gtaogtattt	tgcttaccac	atacagcaat	4380
agtaatggct	ctcagtcctc	gtggatggag	gagcaaatc	gggatgcctg	gggaagcatg	4440
gttctaaaaa	atgttgtagc	tgaaacggat	gaagttggta	aaggtcagat	ccggatgaga	4500
actgtttttg	aacaggccat	tgatcaacgc	tcttcaactg	gtgcctggag	aatgtctctt	4560
tctatttggt	aacctgtctg	caatgaaatt	ttcgatcgct	tgattaaacc	acgctgggag	4620
attagataat	gaagcgtgcg	cctgttattc	caaaacatac	gctcaatact	caaccgggtg	4680
aagatacttc	gttatcgaca	ccagctgccc	cgatgggtgga	ttcgtttaatt	gcgcgcgtag	4740
gagtaatggc	tcgcggtaat	gccattactt	tgccctgatg	tggtcgggat	gtgaagttaa	4800
ctcttgaagt	gctccgggtt	gatagtgttg	agaagacctc	tcgggtatgg	tcaggtaatg	4860
aacgtgacca	ggagctgctt	actgaggacg	cactgcatga	tctcatccct	tcttttctac	4920
tgactgggtc	acagacaccg	gcgttcggtc	gaagagtatc	tggtgtcata	gaaattgccc	4980
atgggagtcg	ccgtcgtaaa	gctgctgcac	ttaccgaaag	tgattatcgt	gttctgggtg	5040
gcgagctgga	tgatgagcag	atggctgcat	tatccagatt	gggtaacgat	tatcgcccaa	5100
caagtgcctt	tgaacgtggt	cagcgttatg	caagccgatt	gcagaatgaa	tttgctggaa	5160
atatttctgc	gctggctgat	gcggaaaata	tttcacgtaa	gattattacc	cgctgtatca	5220
acaccgccaa	attgcctaaa	tcagttgttg	ctcttttttc	tcaccccggt	gaactatctg	5280
cccgtcagg	tgatgcactt	caaaaagcct	ttacagataa	agaggaatta	cttaagcagc	5340
aggcatctaa	ccttcattgag	cagaaaaaag	ctggggtgat	atttgaagct	gaagaagtta	5400
tcactctttt	aacttctgtg	cttaaaacgt	catctgcatac	aagaactagt	ttaagctcac	5460
gacatcagtt	tgctcctgga	gcgacagtat	tgtataaggg	cgataaaatg	gtgcttaacc	5520
tggacaggtc	tcgtgttcca	actgagtgtg	tagagaaaat	tgaggccatt	cttaaggaac	5580
ttgaaaagcc	agcaccttga	tgcgaccacg	ttttagtcta	cgttttatctg	tctttactta	5640
atgtcctttg	ttacaggcca	gaaagcataa	ctggcctgaa	tattctctct	gggcccactg	5700
ttccacttgt	atcgtcggtc	tgataatcag	actgggacca	cgggtcccact	cgtatcgctg	5760

gtctgattat	tagtctggga	ccacgggtccc	actcgtatcg	tcgggtctgat	tattagttctg	5820
ggaccacggt	cccactcgta	tcgctcgggtct	gataatcaga	ctggggaccac	ggtcccactc	5880
gtatcgctcg	tctgattatt	agtctggggac	catgggtccca	ctcgtatcgt	cggctctgatt	5940
attagttctgg	gaccacgggtc	ccactcgtat	cgtcgggtctg	attattagtc	tggaaaccacg	6000
gtcccactcg	tatcgctcgg	ctgattatta	gtctggggacc	acgggtcccac	tcgtatcgtc	6060
gggtctgatta	ttagtctggg	accacgatcc	cactcgtgtt	gtcgggtctga	ttatcgggtct	6120
gggaccacgg	tcccacttgt	attgtcgatc	agactatcag	cgtgagacta	cgattccatc	6180
aatgcctgtc	aagggcaagt	attgacatgt	cgtcgttaacc	tgtagaacgg	agtaacctcg	6240
gtgtgcgggt	gtatgcctgc	tgtggattgc	tgtgtgtcc	tgttatcca	caacattttg	6300
cgcacgggta	tgtggacaaa	atacctgggt	acgttaacca	ggcgtgccc	gcacgttaac	6360
cgggctgcat	ccgatgcaag	tgtgtcgtcg	tcgagtaact	ataacgggtcc	taaggtagcg	6420
aggcgcgcca	tttgggctcg	acattgatta	ttgactagtt	attaatagta	atcaattacg	6480
gggtcattag	ttcatagccc	atatatggag	ttccgcgtta	cataacttac	ggtaaatggc	6540
ccgcttggt	gaccgcccac	cgacccccgc	ccattgacgt	caataatgac	gtatgttccc	6600
atagtaacgc	caatagggac	tttccattga	cgtcaatggg	tggagtattt	acggtaaac	6660
gcccacttgg	cagtacatca	agtgtatcat	atgccaaagta	cgccccctat	tgacgtcaat	6720
gacggtaaat	ggcccgctg	gcattatgcc	cagtacatga	ccttatggga	ctttcctact	6780
tggcagtaca	tctacgtatt	agtcacgtct	attaccatgg	tcgaggtgag	ccccacgttc	6840
tgcttcactc	tccccatctc	ccccccctcc	ccacccccaa	ttttgtattt	atttattttt	6900
taattatttt	gtgcagcgat	gggggcgggg	gggggggggg	cgcgcgccag	gcggggcggg	6960
gcggggcgag	gggcggggag	gggcgaggcg	gagaggtgag	gcggcgacca	atcagagcgg	7020
cgcgtccga	aagtttccct	ttatggcgag	gcggcgggcg	cggcgggcct	ataaaaagcg	7080
aagcgcgcg	cgggcgggag	tcgctgcgtt	gccttcgccc	cgtgccccgc	tccgcgcgcg	7140
ctcgcgcgcg	ccgccccggc	tctgactgac	cgcgttaact	ccacaggtga	gcgggctggg	7200
cggcccttct	cctccgggct	gtaattagcg	cttggtttaa	tgacgggtcg	tttcttttct	7260
gtggctgcgt	gaaagcctta	aagggtccg	ggagggcoct	ttgtgcgggg	gggagcggct	7320
cggggggtgc	gtgcgtgtgt	gtgtgcgtgg	ggagcgccgc	gtgcggcccg	cgctgcccgg	7380
cggctgtgag	cgctgcgggc	gcggcgcggg	gctttgtgag	ctccgcgtgt	gcgcgagggg	7440
agcgcggccg	ggggcggtgc	cccgcgggtgc	gggggggctg	caggggggaa	aaaggctgcg	7500
tgcggggtgt	gtgcgtgggg	gggtgagcag	ggggtgtggg	cgcggcggtc	gggctgtaac	7560
ccccccctgc	acccccctcc	ccgagttgct	gagcacggcc	cggcttcggg	tgcggggctc	7620
cgtacggggc	gtggcgcggg	gctgcgcgtg	ccgggcgggg	ggtggcgggc	ggtgggggtc	7680
ccgggcgggg	cggggcggcc	tcgggcgggg	gagggtcggg	gggagggggc	cggcgccccc	7740
cggagcgccg	gcggctgtcg	aggcgcgggc	agccgcagcc	attgcctttt	atggtaatcg	7800
tgcgagaggg	cgcagggaact	tcctttgtcc	caaactctgt	cggagccgaa	atctgggagg	7860
cgcgcgcgca	ccccctctag	cgggcgcggg	gcgaagcggg	gcggcgccgg	cagggaaggaa	7920
atgggcgggg	agggccttcg	tgcgtgcgcg	cgcgcgcgtc	cccttctccc	tctccagcct	7980
cggggctgtc	cgcgggggga	cggctgcctt	cgggggggac	ggggcgaggc	gggggttcggc	8040
ttctgcgctg	tgaccggcgg	ctctagagcc	tctgctaacc	atgttcatgc	cttcttcttt	8100
ttctacagc	tactgggcaa	cgtgctggtt	attgtgctgt	ctcatcattt	tggcaaaaga	8160
ttcaccatgg	accctgatga	tgttgttgat	tcttctaact	cttttgtgat	ggaaaaacttt	8220
tcttcgtacc	acgggactaa	acctggttat	gtagattcca	ttcaaaaagg	tatacaaaaag	8280
ccaaaatctg	gtacacaagg	aaattatgac	gatgattgga	aagggtttta	tagtaccgac	8340
aataaatacg	acgctgcggg	atactctgta	gataatgaaa	acccgctctc	tggaaaagct	8400
ggaggcggtg	tcaaagtgac	gtatccagga	ctgacgaagg	ttctcgcact	aaaagtggat	8460
aatgcogaaa	ctattaagaa	agagttaggt	ttaagtctca	ctgaaccgtt	gatggagcaa	8520
gtcggaaacg	aagagtttat	caaaaagggtc	ggtgatggtg	cttcgcgtgt	agtgtctcagc	8580
cttcccttcg	ctgaggggag	ttctagcgtt	gaatatatta	ataactggga	acaggcgaaa	8640
gcgttaagcg	tagaacttga	gattaatttt	gaaaccgctg	gaaaacgtgg	ccaagatgcg	8700
atgtatgagt	atatggctca	agcctgtgca	ggaaatcgtg	tcaggcgatc	tttgtgagaa	8760
ttcactcctc	aggtgcaggc	tgcctatcag	aagggtgtgg	ctgggtgtggc	caatgccttg	8820
gctcacaat	accactgaga	tctttttccc	tctgcacaaa	attatgggga	catcatgaag	8880
ccccttgagc	atctgacttc	tggctaataa	aggaaattta	ttttcattgc	aatagtgtgt	8940
tggaaatttt	tgtgtctctc	actcggaagg	acatatggga	gggcaaatca	tttaaaacat	9000
cagaatgagt	atttggttta	gagtttggca	acatatgccc	atatgctggc	tgccatgaac	9060
aaagggttggc	tataaagagg	tcacatgagt	atgaaacagc	cccctgctgt	ccattcctta	9120
ttccatagaa	aagccttgac	ttgaggttag	atttttttta	tattttgttt	tgtgttattt	9180
tttcttttaa	catccctaaa	attttcctta	catgttttac	tagccagatt	tttctctctc	9240
tctgactac	tcccagtcac	agctgtccct	cttctcttat	ggagatccct	cgactgcag	9300
cacaagctcc	ctcgagggag	cttggcgtaa	tcattggtcat	agctgtttcc	tgtgtgaaat	9360
tgttatccgc	tcacaattcc	acacaacata	cgaagccgaa	gcataaagtg	taaagcctgg	9420

ggtgccta	gagtgagcta	actcacatta	attgcgttgc	gctcactgcc	cgctttccag	9480
tcgggaaacc	tgtcgtgcca	gcggatcgat	ccgcatctca	attagtcagc	aaccatagtc	9540
ccgcccctaa	ctccgcccct	cccgccccta	actccgccc	gttccgccc	ttctccgccc	9600
catggctgac	taattttttt	tatttatgca	gaggccgagg	ccgcctcggc	ctctgagcta	9660
ttccagaagt	agtgaggagg	cttttttggg	ggcctaggct	tttgcaaaaa	gctaacttgt	9720
ttattgcagc	ttataatggt	tacaaataaa	gcaatagcat	cacaaatttc	acaaataaag	9780
catttttttc	actgcattct	agttgtggtt	tgtccaaact	catcaatgta	tcttatcatg	9840
tctggatctg	tcgagggccg	gccgagctca	tgcattgcgg	ccgcgtttta	acatttaaat	9900
gtaatacgac	tcactatagg	gcgaggatcc	aagcttagta	ttctatagtg	tcacctaaat	9960
cgtatgtcga	ccggaccggc	ccgggcgcgt	gcttaataaa	tggcaaacag	ctattatggg	10020
tattatgggt	ctcgacattg	attattgact	agtaattcaat	agtaattcaat	tacggggtca	10080
ttagtccata	gcccataatat	ggagttccgc	gttacataac	ttacggtaaa	tggcccgccct	10140
ggctgaccgc	ccaacgaccc	ccgcccattg	acgtcaataa	tgacgtatgt	tcccatagta	10200
acgccaatag	ggactttcca	ttgacgtcaa	tgggtggagt	atttacggta	aactgcccac	10260
ttggcagtac	atcaagtgtg	tcatatgcc	agtacgcccc	ctattgacgt	caatgacggt	10320
aaatggcccg	cctggcatta	tgcccagtag	atgaccttat	gggactttcc	tacttggcag	10380
tacatctacg	tatttagtcat	cgctattacc	atggctcgagg	tgagccccac	gttctgcttc	10440
actctcccca	tctccccccc	ctccccaccc	ccaattttgt	atatttttat	tttttaatta	10500
ttttgtgcag	cgatgggggc	gggggggggg	ggggcgcgcg	ccaggcgggg	cgggggcggg	10560
cgagggggcg	ggcgggcgga	ggcgagagag	tgcggcgcca	gccaatcaga	gcggcgcgct	10620
ccgaaagttt	ccttttatgg	cgagggcgcg	gcggcgcgcg	ccctataaaa	agcgaagcgc	10680
gcggcgggcg	ggagtgcgtg	cgttgccttc	gccccgtgcc	ccgctccgcg	ccgcctcgcg	10740
ccgcccgcgc	cggtcttgac	tgaccgcgtt	actcccacag	gtgagcgggc	gggacggccc	10800
ttctcctccg	ggctgtaatt	agcgtttggt	ttaatgacgg	ctcgtttctt	ttctgtggct	10860
gcgtgaaagc	cttaaagggg	tccgggaggg	ccctttgtgc	gggggggagc	ggctcggggg	10920
gtgcgtgcgt	gtgtgtgtgc	gtggggagcg	ccgcgtgcgg	ccgcgcgtgc	ccggcggtcg	10980
tgagcgctgc	gggcgcggcg	cggggctttg	tgcgtccgc	gtgtgcgcga	ggggagcgcg	11040
gccggggcg	gtgccccgcg	gtgcgggggg	gctgcgaggg	gaacaaaggc	tgcgtgcggg	11100
gtgtgtgcgt	gggggggtga	gcaggggggt	tgggcgcggc	ggtcgggctg	taaccccccc	11160
ctgcaccccc	ctccccgagt	tgctgagcac	ggccccgctt	cggtgtcggg	gctccgtacg	11220
gggcgtggcg	cggggctcgc	cgtgccgggc	ggggggtggc	ggcaggtggg	ggtgccgggc	11280
ggggcggggc	cgccctcggg	cggggagggc	tcgggggagg	ggcgcgggcg	cccccgagc	11340
gccggcggt	gtcgaggcgc	ggcgagccgc	agccattgcc	ttttatggta	atcggtcgag	11400
agggcgccag	gacttccctt	gtcccaaatc	tgtgcggagc	cgaaatctgg	gaggcgccgc	11460
cgcacccctc	ctagcgggcg	cggggcgaag	cggtgcggcg	ccggcaggaa	ggaaatgggc	11520
ggggagggcc	ttcgtgcgtc	gccgcgcgcg	cgcccccttc	tccctctcca	gcctcggggc	11580
tgtccgcggg	gggacggctg	ccttcggggg	ggacggggca	gggcgggggt	cggtctctgg	11640
cgtgtgaccg	gcggctctag	agcctctgct	aacctgttcc	atgccttctt	ctttttccta	11700
cagctcctgg	gcaacgtgct	ggttattgtg	ctgtctcatc	attttgcaa	agaattcacc	11760
atggaccctg	atgatgttgt	tgattcttct	aaatcttttg	tgatggaaaa	cttttcttcg	11820
taccacggga	ctaaacctgg	ttatgtagat	tccattcaaa	aaggtataca	aaagccaaaa	11880
tctggtacac	aaggaaatta	tgacgatgat	tggaaagggt	tttatagtag	cgacaataaa	11940
tacgacgtgc	cgggatactc	tgtagataat	gaaaaccgcg	tctctggaaa	agctggaggc	12000
gtggtcaaat	tgacgtatcc	aggactgacg	aaggttctcg	cactaaaagt	ggataatgcc	12060
gaaactatta	agaaagagtt	aggtttaagt	ctcactgaac	cgttgatgga	gcaagtcgga	12120
acggaagagt	ttatcaaaag	gttcggtgat	ggtgcttcgc	gtgtagtgct	cagccttccc	12180
ttcgctgagg	ggagttctag	cgttgaatat	attaataact	gggaacaggc	gaaagcggtta	12240
agcgtagaac	ttgagattaa	ttttgaaacc	cgtggaaaac	gtggccaaga	tgcatgtgat	12300
gagtatatgg	ctcaagcctg	tgcaggaaat	cgtgtcaggc	gatctttgtg	agaattcact	12360
cctcaggtgc	aggctgccta	tcagaagggt	gtggctgggt	tggccaatgc	cctggctcac	12420
aaataccact	gagatctttt	tccctctgcc	aaaaattatg	gggacatcat	gaagcccctt	12480
gagcatctga	cttctggcta	ataaaggaaa	tttattttca	ttgcaatagt	gtgttggaat	12540
tttttgtgtc	tctcactcgg	aaggacatat	gggagggcaa	atcattttaa	acatcagaat	12600
gagtatttgg	tttagagttt	ggcaacatat	gcccataatg	tggctgccat	gaacaaagg	12660
tggctataaa	gaggtcatca	gtatatgaaa	cagccccctg	ctgtccattc	cttattccat	12720
agaaaagcct	tgacttgagg	ttagattttt	tttatatttt	gttttgtgtt	atttttttct	12780
ttaacatccc	taaaattttc	cttacatgtt	ttactagcca	gatttttctt	cctctcctga	12840
ctactcccag	tcatagctgt	ccctcttctc	ttatggagat	ccctcgacct	gcagcacaag	12900
ctccctcgag	ggagcttggc	gtaatcatgg	tcatagctgt	ttcctgtgtg	aaattgttat	12960
ccgctcacia	ttccacacia	catacgagcc	ggaagcataa	agtgtaaagc	ctggggtgcc	13020
taatgagtga	gctaactcac	attaattgcg	ttgcgctcac	tgcccgcttt	ccagtcggga	13080

```

aacctgtcgt gccagcggat cgatccgcat ctcaattagt cagcaaccat agtcccgcgc 13140
ctaactccgc ccatcccgcc cctaactccg ccaggttccg cccattctcc gcccctggc 13200
tgactaattt tttttattta tgcagaggcc gagggcgccct cggcctctga gctattccag 13260
aagtagtgag gaggtttttt tggaggccta ggcttttgca aaaagctaac ttgtttattg 13320
cagcttataa tggttacaaa taaagcaata gcatacaaaa tttcacaaat aaagcatttt 13380
tttcactgca ttctagttgt ggtttgtcca aactcatcaa tgtatcttat catgtctgga 13440
tcaaattctcg actcgctacc ttaggaccgt tatagtta 13478

```

<210> 2

<211> 84

<212> ADN

<213> Séquence artificielle

<220>

<223> Description de la séquence artificielle: Séquence
contenant des sites de restrictions

<400> 2

ctcgagtaac tataacggtc ctaaggtagc gagggcggcc atcgatgtcg actcgctacc 60

ttaggaccgt tatagttact cgag

84

<210> 3

<211> 22

<212> ADN

<213> Séquence artificielle

<220>

<223> Description de la séquence artificielle: Séquence
contenant des sites de restrictions

<400> 3

ggcgcgccat ttaaattctcg ag

22

<210> 4

<211> 186

<212> ADN

<213> Séquence artificielle

<220>

<223> Description de la séquence artificielle: Séquence
contenant des sites de restrictions

<400> 4

tcgagggccg gccgagctca tgcattgcgg ccgcgtttta acatttaaata gtaatacgac 60

tcactatagg gcgaggatcc aagcttagta ttctatagtg tcacctaaat cgtatgtoga 120

ccggaccggc ccggggcgcat gcttaattaa tggcaaacag ctattatggg tattatgggt 180

ctcgag

186

(12) DEMANDE INTERNATIONALE PUBLIÉE EN VERTU DU TRAITÉ DE COOPÉRATION
EN MATIÈRE DE BREVETS (PCT)

(19) Organisation Mondiale de la Propriété
Intellectuelle
Bureau international



(43) Date de la publication internationale
5 décembre 2002 (05.12.2002)

PCT

(10) Numéro de publication internationale
WO 02/097100 A3

(51) Classification internationale des brevets⁷ : C12N 15/79

(21) Numéro de la demande internationale :
PCT/FR02/01782

(22) Date de dépôt international : 28 mai 2002 (28.05.2002)

(25) Langue de dépôt : français

(26) Langue de publication : français

(30) Données relatives à la priorité :
0106921 28 mai 2001 (28.05.2001) FR

(71) Déposant (pour tous les États désignés sauf US) :
GENOWAY [FR/FR]; 181, Avenue Jean-Jaurès, Immeu-
ble ChâteauBriand, F-69007 Lyon (FR).

(72) Inventeurs; et

(75) Inventeurs/Déposants (pour US seulement) : LAPIZE-
GAUTHEY, Christine [FR/FR]; 4, rue Latreille, F-38200
Vienne (FR). FRAICHARD, Alexandre [FR/FR]; 144,
avenue des Etats-Unis, F-78000 Versailles (FR).

(74) Mandataires : MARTIN, Jean-Jacques etc.; Cabinet
Regimbeau, 26, rue de Chazelles, F-75847 Paris Cedex 17
(FR).

(81) États désignés (national) : AE, AG, AL, AM, AT, AU, AZ,
BA, BB, BG, BR, BY, BZ, CA, CH, CN, CO, CR, CU, CZ,

DE, DK, DM, DZ, EC, EE, ES, FI, GB, GD, GE, GH, GM,
HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK,
LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX,
MZ, NO, NZ, OM, PH, PL, PT, RO, RU, SD, SE, SG, SI,
SK, SL, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VN,
YU, ZA, ZM, ZW.

(84) États désignés (régional) : brevet ARIPO (GH, GM, KE,
LS, MW, MZ, SD, SL, SZ, TZ, UG, ZM, ZW), brevet
eurasien (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), brevet
européen (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR,
IE, IT, LU, MC, NL, PT, SE, TR), brevet OAPI (BF, BJ,
CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN,
TD, TG).

Déclaration en vertu de la règle 4.17 :

— relative à la qualité d'inventeur (règle 4.17.iv)) pour US
seulement

Publiée :

— avec rapport de recherche internationale

(88) Date de publication du rapport de recherche
internationale: 20 mars 2003

En ce qui concerne les codes à deux lettres et autres abrégia-
tions, se référer aux "Notes explicatives relatives aux codes et
abrégiactions" figurant au début de chaque numéro ordinaire de
la Gazette du PCT.

(54) Title: CLONING VECTORS FOR HOMOLOGOUS RECOMBINATION AND METHOD USING SAME

(54) Titre : VECTEURS DE CLONAGE POUR RECOMBINAISON HOMOLOGUE ET PROCÉDE LES UTILISANT

(57) Abstract: The invention concerns novel vectors for use both in creating genomic DNA banks and as vectors for carrying out homologous recombination reactions in host cells, in particular for improving selection of said homologous recombination events, and decreasing the time for obtaining a final vector for performing the homologous recombination reaction. The invention also concerns a method using said vectors.

(57) Abrégé : La présente invention se rapporte à des nouveaux vecteurs utilisables à la fois pour créer des banques d'ADN gé-
nomiques et en tant que vecteurs pour la réalisation de réactions de recombinaison homologue dans des cellules hôtes, permettant
notamment d'améliorer la sélection desdits événements de recombinaison homologue, et de diminuer le temps d'obtention du vec-
teur final pour effectuer la réaction de recombinaison homologue. L'invention a également pour objet un procédé mettant en oeuvre
ces vecteurs.



WO 02/097100 A3

INTERNATIONAL SEARCH REPORT

International Application No

PCT/FR 02/01782

A. CLASSIFICATION OF SUBJECT MATTER
IPC 7 C12N15/79

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)
IPC 7 C12N

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practical, search terms used)

EPO-Internal, WPI Data, BIOSIS

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category *	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	MAHILLON J ET AL: "New ultrarare restriction site-carrying transposons for bacterial genomics" GENE: AN INTERNATIONAL JOURNAL ON GENES AND GENOMES, ELSEVIER SCIENCE PUBLISHERS, BARKING, GB, vol. 187, no. 2, 18 March 1997 (1997-03-18), pages 273-279, XP004093291 ISSN: 0378-1119 the whole document	1-4
A	US 5 874 259 A (SZYBALSKI WACLAW) 23 February 1999 (1999-02-23) abstract; figure 1 column 4, line 31 -column 8, line 23 Seq.Id.No.1 ----- -/-	12

☒ Further documents are listed in the continuation of box C.

☒ Patent family members are listed in annex.

* Special categories of cited documents:

- *A* document defining the general state of the art which is not considered to be of particular relevance
- *E* earlier document but published on or after the international filing date
- *L* document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)
- *O* document referring to an oral disclosure, use, exhibition or other means
- *P* document published prior to the international filing date but later than the priority date claimed

- *T* later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention
- *X* document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone
- *Y* document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art
- *Z* document member of the same patent family

Date of the actual completion of the international search

20 December 2002

Date of mailing of the international search report

07/01/2003

Name and mailing address of the ISA

European Patent Office, P.B. 5818 Patentlaan 2
NL - 2280 HV Rijswijk
Tel. (+31-70) 340-2040, Tx. 31 651 epo nl,
Fax: (+31-70) 340-3016

Authorized officer

Panzica, G

INTERNATIONAL SEARCH REPORT

International Application No
PCT/FR 02/01782

C.(Continuation) DOCUMENTS CONSIDERED TO BE RELEVANT

Category *	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	WO 00 26386 A (UNIV GEORGIA) 11 May 2000 (2000-05-11) abstract figure 1 exemples	1-4,6-11
A	WO 98 58067 A (EDGE BIOSYSTEMS INC ;QIAN SU WEN (US); SEED JOHN (US); CANDAU REYE) 23 December 1998 (1998-12-23) abstract figure 1	1-4,6-11
A	US 5 672 510 A (ANDERSON W FRENCH ET AL) 30 September 1997 (1997-09-30) abstract figure 1	1-4,6-11

INTERNATIONAL SEARCH REPORT

tional Application No

PCT/FR 02/01782

Patent document cited in search report		Publication date		Patent family member(s)	Publication date
US 5874259	A	23-02-1999	AU	1463099 A	15-06-1999
			CA	2308608 A1	03-06-1999
			WO	9927118 A1	03-06-1999
WO 0026386	A	11-05-2000	US	6096523 A	01-08-2000
			AU	1335900 A	22-05-2000
			WO	0026386 A1	11-05-2000
WO 9858067	A	23-12-1998	AU	7974498 A	04-01-1999
			EP	1012314 A1	28-06-2000
			WO	9858067 A1	23-12-1998
US 5672510	A	30-09-1997	CA	2034533 A1	20-07-1991
			EP	0511311 A1	04-11-1992
			JP	5503852 T	24-06-1993
			WO	9110728 A1	25-07-1991

RAPPORT DE RECHERCHE INTERNATIONALE

e Internationale No

PCT/FR 02/01782

A. CLASSEMENT DE L'OBJET DE LA DEMANDE
CIB 7 C12N15/79

Selon la classification internationale des brevets (CIB) ou à la fois selon la classification nationale et la CIB

B. DOMAINES SUR LESQUELS LA RECHERCHE A PORTE

Documentation minimale consultée (système de classification suivi des symboles de classement)

CIB 7 C12N

Documentation consultée autre que la documentation minimale dans la mesure où ces documents relèvent des domaines sur lesquels a porté la recherche

Base de données électronique consultée au cours de la recherche internationale (nom de la base de données, et si réalisable, termes de recherche utilisés)

EPO-Internal, WPI Data, BIOSIS

C. DOCUMENTS CONSIDERES COMME PERTINENTS

Catégorie *	Identification des documents cités, avec, le cas échéant, l'indication des passages pertinents	no. des revendications visées
A	MAHILLON J ET AL: "New ultrarare restriction site-carrying transposons for bacterial genomics" GENE: AN INTERNATIONAL JOURNAL ON GENES AND GENOMES, ELSEVIER SCIENCE PUBLISHERS, BARKING, GB, vol. 187, no. 2, 18 mars 1997 (1997-03-18), pages 273-279, XP004093291 ISSN: 0378-1119 le document en entier ---	1-4
A	US 5 874 259 A (SZYBALSKI WACLAW) 23 février 1999 (1999-02-23) abrégé; figure 1 colonne 4, ligne 31 -colonne 8, ligne 23 Seq.Id.No.1 --- -/--	12



Voir la suite du cadre C pour la fin de la liste des documents



Les documents de familles de brevets sont indiqués en annexe

* Catégories spéciales de documents cités:

- *A* document définissant l'état général de la technique, non considéré comme particulièrement pertinent
- *E* document antérieur, mais publié à la date de dépôt international ou après cette date
- *L* document pouvant jeter un doute sur une revendication de priorité ou cité pour déterminer la date de publication d'une autre citation ou pour une raison spéciale (telle qu'indiquée)
- *O* document se référant à une divulgation orale, à un usage, à une exposition ou tous autres moyens
- *P* document publié avant la date de dépôt international, mais postérieurement à la date de priorité revendiquée

T document ultérieur publié après la date de dépôt international ou la date de priorité et n'appartenant pas à l'état de la technique pertinent, mais cité pour comprendre le principe ou la théorie constituant la base de l'invention

X document particulièrement pertinent; l'invention revendiquée ne peut être considérée comme nouvelle ou comme impliquant une activité inventive par rapport au document considéré isolément

Y document particulièrement pertinent; l'invention revendiquée ne peut être considérée comme impliquant une activité inventive lorsque le document est associé à un ou plusieurs autres documents de même nature, cette combinaison étant évidente pour une personne du métier

Z document qui fait partie de la même famille de brevets

Date à laquelle la recherche internationale a été effectivement achevée

20 décembre 2002

Date d'expédition du présent rapport de recherche internationale

07/01/2003

Nom et adresse postale de l'administration chargée de la recherche internationale

Office Européen des Brevets, P.B. 5818 Patentlaan 2
NL - 2280 HV Rijswijk
Tel. (+31-70) 340-2040, Tx. 31 651 epo nl,
Fax: (+31-70) 340-3016

Fonctionnaire autorisé

Panzica, G

C.(suite) DOCUMENTS CONSIDERES COMME PERTINENTS		
Catégorie	Identification des documents cités, avec, le cas échéant, l'indication des passages pertinents	no. des revendications visées
A	WO 00 26386 A (UNIV GEORGIA) 11 mai 2000 (2000-05-11) abrégé figure 1 exemples ----	1-4,6-11
A	WO 98 58067 A (EDGE BIOSYSTEMS INC ; QIAN SU WEN (US); SEED JOHN (US); CANDAU REYE) 23 décembre 1998 (1998-12-23) abrégé figure 1 -----	1-4,6-11
A	US 5 672 510 A (ANDERSON W FRENCH ET AL) 30 septembre 1997 (1997-09-30) abrégé figure 1 -----	1-4,6-11

RAPPORT DE RECHERCHE INTERNATIONALE

de Internationale No
PCT/FR 02/01782

Document brevet cité au rapport de recherche		Date de publication	Membre(s) de la famille de brevet(s)		Date de publication
US 5874259	A	23-02-1999	AU	1463099 A	15-06-1999
			CA	2308608 A1	03-06-1999
			WO	9927118 A1	03-06-1999
WO 0026386	A	11-05-2000	US	6096523 A	01-08-2000
			AU	1335900 A	22-05-2000
			WO	0026386 A1	11-05-2000
WO 9858067	A	23-12-1998	AU	7974498 A	04-01-1999
			EP	1012314 A1	28-06-2000
			WO	9858067 A1	23-12-1998
US 5672510	A	30-09-1997	CA	2034533 A1	20-07-1991
			EP	0511311 A1	04-11-1992
			JP	5503852 T	24-06-1993
			WO	9110728 A1	25-07-1991